NPS ARCHIVE
1997
EMMONS, N.

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

MATHEMATICAL MODELING USING MICROSOFT EXCEL

by

Nelson L. Emmons, Jr.

June 1997

Thesis Advisor: Maurice D. Weir

Approved for public release; distribution is unlimited.

Thesis
E436334

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1997	Master's Thesis

4. TITLE AND SUBTITLE
MATHEMATICAL MODELING USING MICROSOFT EXCEL

5. FUNDING NUMBERS

6. AUTHOR(S)
Emmons, Nelson L., Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The entry into higher mathematics begins with calculus. Rarely, however, does the calculus student recognize the full power and applications for the mathematical concepts and tools that are taught. Frank R. Giordano, Maurice D. Weir, and William P. Fox produced A First Course in Mathematical Modeling, a unique text designed to address this shortcoming and teach the student how to identify, formulate, and interpret the real world in mathematical terms. Mathematical modeling is the application of mathematics to explain or predict real world behavior. Often real world data are collected and used to verify or validate (and sometimes formulate) a hypothetical model or scenario. Inevitably, in such situations, it is desirable and necessary to have computational support available to analyze the large amounts of data. Certainly this eliminates the tedious and inefficient hand calculations necessary to validate and apply the model (assuming the calculations can even be reasonably done by hand).

The primary purpose of *Mathematical Modeling Using Microsoft Excel* is to provide instructions and examples for using the spreadsheet program Microsoft Excel to support a wide range of mathematical modeling applications. Microsoft Excel is a powerful spreadsheet program which allows the user to organize numerical data into an easy-to-follow on-screen grid of columns and rows. Our version of Excel is based on Microsoft Windows. In this text, it is not the intent to teach mathematical modeling, but rather to provide computer support for most of the modeling topics covered in A First Course in Mathematical Modeling. The examples given here are support that text as well.

14. SUBJECT TERMS
Discrete Dynamical Systems, Proportionality, Model Fitting, Monte Carlo Simulation, Linear Programming

15. NUMBER OF PAGES
156

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

Approved for public release; distribution is unlimited.

MATHEMATICAL MODELING USING MICROSOFT EXCEL

Nelson L. Emmons, Jr.
Captain, United States Army
B.S., United States Military Academy, 1989

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN APPLIED MATHEMATICS

from the

NAVAL POSTGRADUATE SCHOOL
June 1997

NPS ARCHIVE
1997.06
EMMONS, N.

Thesis
E7363334
C.2

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

ABSTRACT

The entry into higher mathematics begins with calculus. Rarely, however, does the calculus student recognize the full power and applications for the mathematical concepts and tools that are taught. Frank R. Giordano, Maurice D. Weir, and William P. Fox produced A First Course in Mathematical Modeling, a unique text designed to address this shortcoming and teach the student how to identify, formulate, and interpret the real world in mathematical terms. Mathematical modeling is the application of mathematics to explain or predict real world behavior. Often real world data are collected and used to verify or validate (and sometimes formulate) a hypothetical model or scenario. Inevitably, in such situations, it is desirable and necessary to have computational support available to analyze the large amounts of data. Certainly this eliminates the tedious and inefficient hand calculations necessary to validate and apply the model (assuming the calculations can even be reasonably done by hand).

The primary purpose of *Mathematical Modeling Using Microsoft Excel* is to provide instructions and examples for using the spreadsheet program Microsoft Excel to support a wide range of mathematical modeling applications. Microsoft Excel is a powerful spreadsheet program which allows the user to organize numerical data into an easy-to-follow on-screen grid of columns and rows. Our version of Excel is based on Microsoft Windows. In this text, it is not the intent to teach mathematical modeling, but rather to provide computer support for most of the modeling topics covered in A First Course in Mathematical Modeling. The examples given here support that text as well.

TABLE OF CONTENTS

I. AN INTRODUCTION TO MICROSOFT EXCEL 1

A. BACKGROUND AND PURPOSE 1

B. THE STRUCTURE OF MICROSOFT EXCEL 2

C. BASIC COMMANDS USED IN EXCEL 5

1. How to Start a New Workbook 6

2. How to Save, Open, and Close a Workbook 7

3. How to Print a Worksheet 8

4. How to Move through a Worksheet 9

D. ENTERING DATA 10

E. ENTERING FORMULAS 12

1. Use of Arithmetic and Relational Operators 13

2. Formula Error Values 14

3. Use of Cell References 14

F. EDITING A WORKSHEET 16

1. Modifying the Contents of a Cell 16

2. Erasing Data from a Worksheet 17

II. MODELING WITH DISCRETE DYNAMICAL SYSTEMS 19

A. AN EXAMPLE OF FORMULATING DATA 19

Example 2.1: The Arms Race 19

B. CREATING CHARTS FROM WORKSHEET DATA 23

1. How to Create a Chart on a Chart Sheet 25

2. How to Create a Chart on a Worksheet (Embedded) 26

a. An Example of Creating a Chart 27

Example 2.2 (a): The Arms Race Revisited 27
3. How to Modify a Chart 29
Example 2.2 (b): The Arms Race Revisited (continued) 31
4. How to Add and Delete Chart Data 32
Example 2.2 (c): The Arms Race Revisited (continued) 33

C. MODELING CHANGE WITH DIFFERENCE EQUATIONS 34
Example 2.3: Mortgaging a Home 34

D. APPROXIMATING CHANGE WITH DIFFERENCE EQUATIONS 39
Example 2.4: Growth of a Yeast Culture 39

E. NUMERICAL SOLUTIONS 45
Example 2.5: Prescription for Digoxin 45

F. SYSTEMS OF DIFFERENCE EQUATIONS 48
Example 2.6: Competitive Hunter Models -- Spotted Owls and Hawks 48

III. MODELING USING PROPORTIONALITY 53

A. ADVANCED SKILLS FOR BUILDING WORKSHEETS 53
1. Cell Reference Operators 54
2. Built-in Functions 54
3. Use of the Function Wizard 56

B. VEHICULAR STOPPING DISTANCE 57
Example 3.1: Vehicle Stopping Distance 57
1. Adding a Trendline to a Data Series 60
2. Adding Error Bars to a Data Series 61

C. A BASS FISHING DERBY 64
Example 3.2: A Bass Fishing Derby 65

IV. MODEL FITTING AND EXPERIMENTAL MODELING 71

A. MODEL FITTING 71
1. Least-Squares Curve Fitting Using Excel 72
Example 4.1(a): Vehicular Stopping Distance Revisited 73
2. Choosing a Best Model 81
Example 4.1(b): Vehicular Stopping Distance Revisited (continued) 81

B. EXPERIMENTAL MODELING 84
1. High-Order Polynomial Models 85
Example 4.2: Elapsed Time of a Tape Recorder 85
2. Smoothing: Low-Order Polynomial Models 88
a. Divided Differences 89
Example 4.3: Elapsed Time of a Tape Recorder Revisited 89
b. Observations on Difference Tables 93
Example 4.4: Growth of a Yeast Culture Revisited 93
3. Cubic Spline Models 98
Example 4.5: Vehicle Stopping Distance Revisited Again 99

V. SIMULATION MODELING (MONTE CARLO) 101

A. RANDOM NUMBER GENERATION USING EXCEL 101

B. SIMULATING DETERMINISTIC BEHAVIOR 102
Example 5.1: Area Under a Curve 102
Example 5.2: Volume Under a Surface 105

C. SIMULATING PROBABILISTIC BEHAVIOR 107
1. How to Use the Excel Histogram 108
Example 5.3: A Fair Coin 108
Example 5.4: Roll of a Fair Die 110
Example 5.5: Roll of an Unfair Die 112

VI. LINEAR PROGRAMMING 115

A. LINEAR PROGRAMMING: THE SIMPLEX METHOD 115

B. USE OF THE EXCEL SOLVER 116
1. Using the Excel Solver 117
Example 6.1: The Carpenter's Problem 119
Example 6.2: The Carpenter's Problem Revisited 122

C. LINEAR PROGRAMMING: SENSITIVITY ANALYSIS 125
1. Using the Solver Sensitivity Report 126
Example 6.3: The Carpenter's Problem Revisited Again 126

VII. CONCLUSION 129

APPENDIX A. RETRIEVING AND ENTERING DATA FROM OTHER FILES 131

APPENDIX B. NUMERICAL DATA FOR MORTGAGING A HOME 135

APPENDIX C. NUMERICAL SOLUTIONS TO COMPETITIVE HUNTER MODELS 139

LIST OF REFERENCES 143

INITIAL DISTRIBUTION LIST 145

I. AN INTRODUCTION TO MICROSOFT EXCEL

A. BACKGROUND AND PURPOSE

Mathematical modeling is the application of mathematics to explain or predict real world behavior. Often real world data are used to verify or validate a hypothetical model or scenario. The data requirements can be immense when dealing with real world problems. Typically, in such situations, it is desirable to have computational support available to analyze the large amounts of data, therefore eliminating the tedious and inefficient hand calculations necessary to validate and apply the model. The primary purpose of this text is to provide instructions and examples for using the spreadsheet program Microsoft Excel to support the teaching of a wide range of mathematical modeling applications.

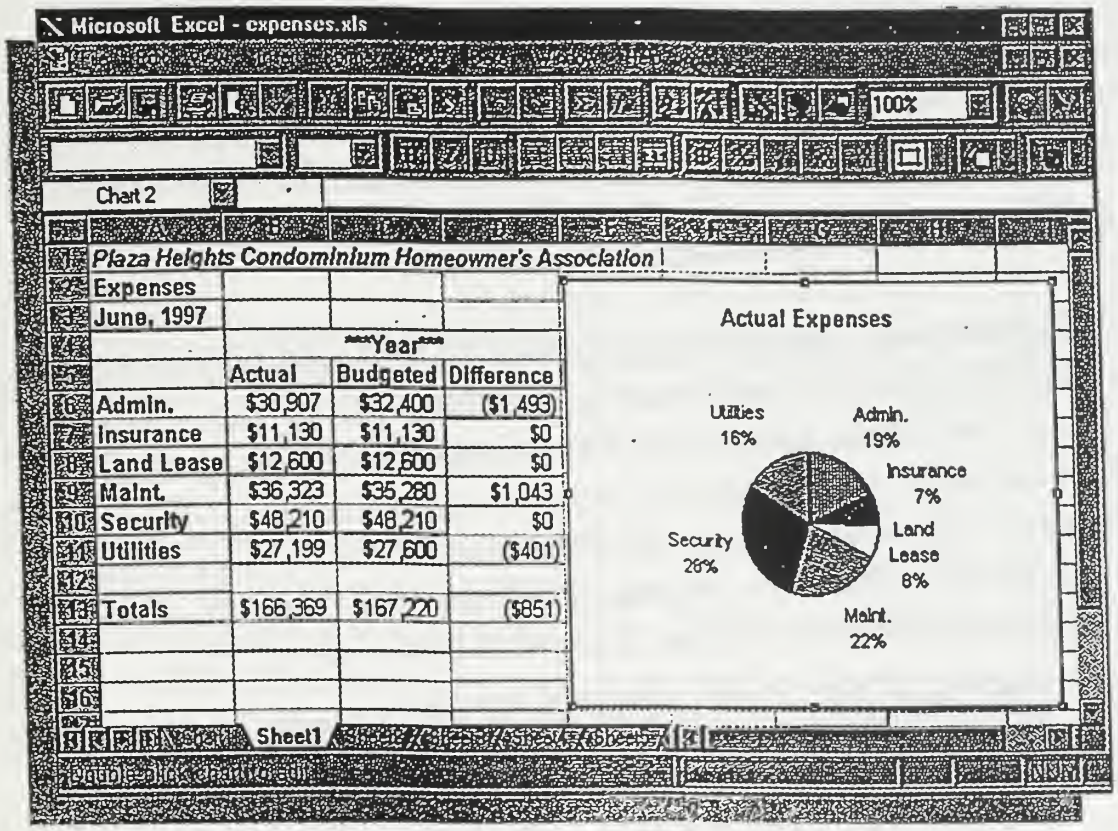


Figure 1-A.1: An Example of a Microsoft Excel Spreadsheet

Microsoft Excel is a powerful spreadsheet program that allows the user to organize numerical data into an easy-to-follow on-screen grid of columns and rows. Excel is just one spreadsheet program among many (e.g.,

Lotus 1-2-3 and Quattro Pro). Microsoft Excel is especially popular among users because of its comfortable "look and feel" and its convenient, versatile features. Excel is based on Microsoft Windows, a program that helps in managing the computing environment and in running other programs. Conveniently, this means that Excel looks and operates similarly to that of many other Windows-based programs (see Figure 1-A.1). A complete description of Microsoft Excel is given in the Microsoft Excel User's Guide. Several versions of Microsoft Excel, as well as of Microsoft Windows, exist. This thesis describes Excel Version 5.0 using Windows Version 3.1. With a different version, one would use the Microsoft Excel User's Guide or consult the installation personnel to determine the necessary adjustments (which should be minor).

In this text, not all of the many standard applications of Microsoft Excel are addressed. Instead, the goal is to present only those Excel commands required to solve specific modeling applications. Each chapter begins with a description of the new Excel commands, a definition of their format commands, and an illustration of their application in short examples. **Boldface** statements are used to define a command format and to enter data, for easy reference. *Italics* represents the introduction of new terminology, and it highlights variables and symbols in the text. Longer modeling applications that illustrate the use of the commands described in that chapter are presented at the end of each chapter. In this text, it is not the intent to teach mathematical modeling, but to provide computer support for the modeling topics covered in A First Course In Mathematical Modeling by Frank Giordano, Maurice Weir, and William Fox. The examples given here are designed in support of that text as well.

B. THE STRUCTURE OF MICROSOFT EXCEL

Microsoft Excel is a spreadsheet program that supplies an on-screen grid of columns and rows into which the user enters data. To start Excel from the Windows menu, double-click the Excel Icon, which loads the Excel program and a blank document (see Figure 1-B.2). No more than three Microsoft Excel programs can be running at one time. The blank Excel document is referred to as a *workbook*, and is the basic file in Excel. As in most Windows programs, more than one workbook can be opened at a time; however, only one of these programs can be active at a time and it is distinguished by a highlighted Title bar. Many of the standard Windows controls and components can be found in Excel's workbook toolbar (see Figure 1-B.1).

Figure 1-B.1: Standard Toolbar with the Formatting Toolbar Underneath

There is the Title bar at the top, with its Maximize, Minimize, and Close buttons in the far right corner. The Title bar will have the words "Microsoft Excel" centered in the middle. Beneath this bar is the Menu bar with the commands File, Edit, View, and so forth. On the far right end of this Menu bar is also an additional Max-Min

button for the work area. If the work area is already minimized, then the button will not appear. Beneath the Menu bar are two toolbars, which offer a variety of icons (such as Save, Print, Copy, Paste, etc.) for performing common Excel procedures (see Figure 1-B.1). Many of the Toolbar icons are equivalent to commands on the Menu bar.

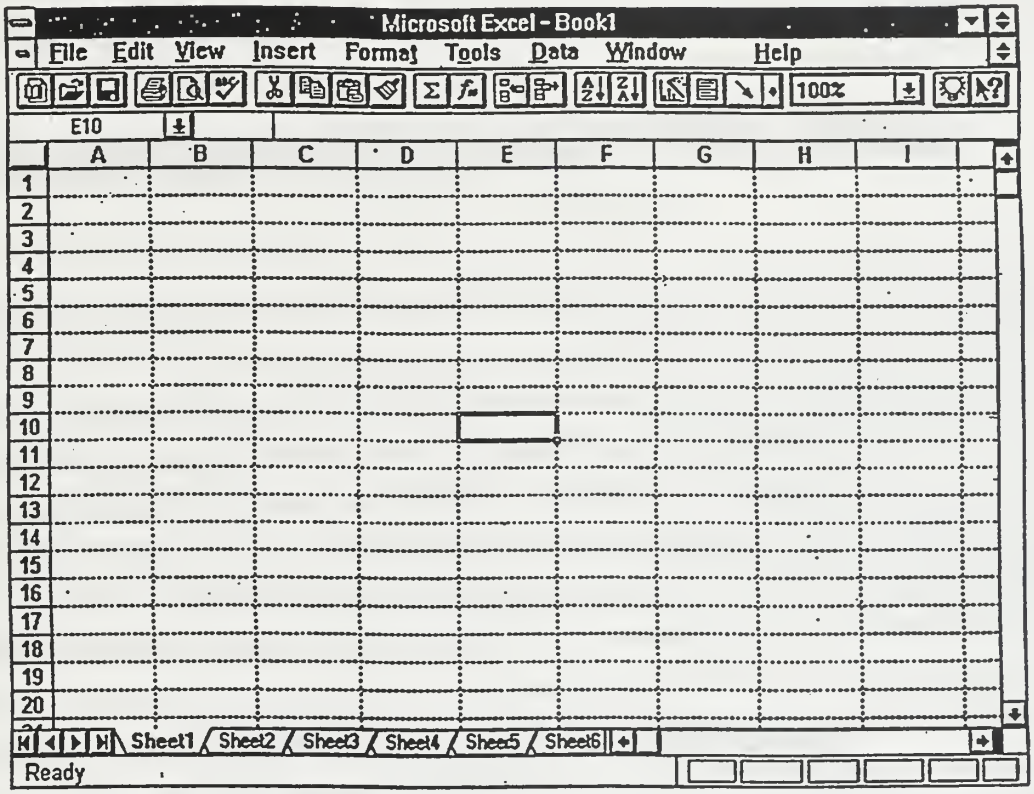

Microsoft Excel - Book1
File Edit View Insert Format Tools Data Window Help
E10
A B C D E F G H I
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Sheet1 Sheet2 Sheet3 Sheet4 Sheet5 Sheet6
Ready

Figure 1-B.2: Blank Microsoft Excel Workbook Document

Below these toolbars is the main body of the workbook or document. It consists of a spreadsheet with rows and columns accompanied with horizontal and vertical scroll bars to help navigate through these rows and columns. This main body or window has its own Title bar with window controls. This is also the location of the file name. File names are automatically given by Excel as Book1, Book2, etc. If the main body is maximized, then the Title bar will be hidden from view and the file name will appear on the Title bar for the Excel program itself (e.g., Microsoft Excel -- Book1). The basic work area in the Excel workbook is called the *spreadsheet*, or *worksheet*. Data and formulas can be entered here. The main body of the workbook contains sixteen (16) worksheets. Each worksheet in a workbook is given the name Sheet1, Sheet2, etc. located at the bottom of the worksheet on its *sheet tab*. To select a sheet, just click on its tab (see Figure 1-B.3).

Figure 1-B.3: Example of the Sheet Tabs

At the bottom of the program window is the *Status bar* which offers various messages (see Figure 1-B.4). The indicator on the left side of the Status bar shows the current state of the program (e.g., it displays "*Ready*" when it is ready to perform any action, or "*Edit*" when a cell is being edited).

Figure 1-B.4: Example of the Status Bar

The right side of the Status bar displays the current state of several toggle keys, including "*CAPS*" when the Caps Lock is on, and "*NUM*" when the Num Lock is on.

Each worksheet in Excel is made up of rows and columns. The rows are identified by numbers, labeled down the left side of the sheet in the heading area. Although only 15 rows can be seen on the screen, there are 65,536 rows in each worksheet. The columns are identified by letters, labeled across the top of the sheet in the heading area. Although only columns A through J can be seen, there are 256 columns in each worksheet. Column Z is followed by columns AA, AB, AC, column AZ is followed by BA, BB, BC, and so on.

At the intersection of each row and column is a *cell*. The cells are where the data (including text, numbers, and formulas) are entered. A particular cell is referred to by its *reference* or *address*, which is a column letter and row number. For example, the cell where column B and row 5 intersect is called cell B5. Generally, one selects the cell or cells to work with, and then the data is entered or a command is selected. Selected cells appear highlighted on the screen. The *active cell* is the cell a user is currently working in, and it is shown by a heavy border. The active cell's address or reference and data information also appear simultaneously in the *Formula bar* just above the worksheet's column lettering (see Figure 1-B.5).

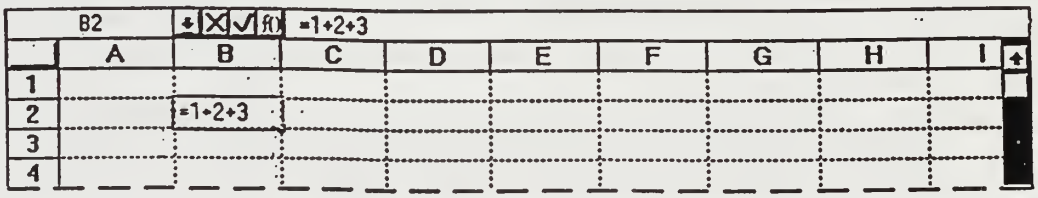


Figure 1-B.5: Example of the Formula Bar (located at the top of the figure)

The Formula bar is located beneath all the toolbars and just above the workbook window. The Formula bar serves two important functions in Excel. First, the Formula bar displays whatever is inside the active cell. When a cell is active or highlighted, although the cell displays a number which is the result of a formula, the actual

formula for that cell is displayed in the Formula bar. Therefore, the second function the Formula bar serves is to enable the user to enter and revise data for the active cell. To the left of the Formula bar is the *name box* which displays the currently active cell's reference/address or name. If the active cell is B2, then the reference B2 will appear in the name box along with the data or formula entered for that cell in the Formula bar.

Finally, hiding behind all the cells, worksheets, and workbooks is Excel's powerful calculating engine that can update thousands of formulas in less time than it takes to type them in. Excel keeps track of all changes made to the worksheet and updates the formulas accordingly.

C. BASIC COMMANDS USED IN EXCEL

Excel offers several ways for the user to interact with the program. The mouse is the most convenient way to issue commands and manipulate items on the screen. The mouse pointer takes on the shape of a plus sign when it is over a cell. Although the keyboard must be used to enter text or numbers, it is not the best way for moving through a Windows-based program such as Microsoft Excel. Almost all window-based mouse techniques apply to Excel (e.g., dragging, double-clicking, use of scroll bars, changing window size, etc.).

Like all Windows-based programs, Microsoft Excel includes a Menu bar which can pull down a list of possible choices under each menu. Many commands can be carried out by just selecting them from their menus. Some commands need additional information before proceeding with their execution. Therefore an additional window may appear called a *dialog box* (see Figure 1-C.1).

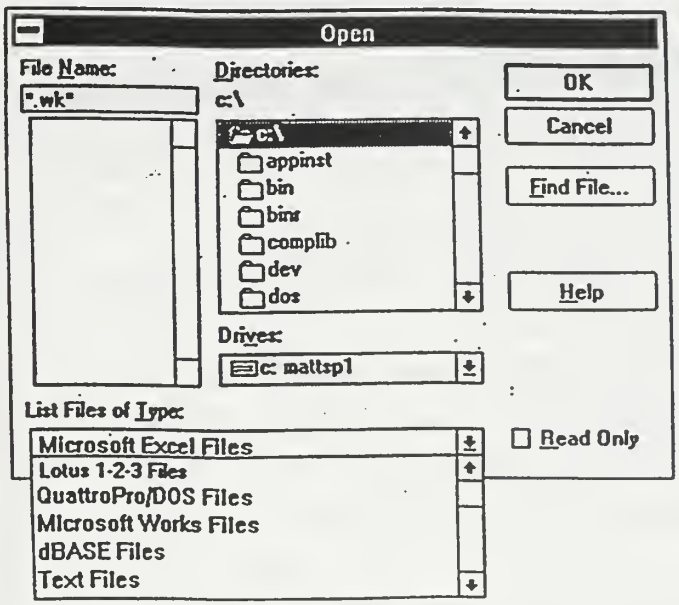


Figure 1-C.1: Example of the Open Dialog Box
(from the **OPEN** command under the **FILE** menu)

The dialog box includes selecting a particular option or choosing from a prescribed list. Once all selections are made, click the **OK** button to execute the command (which also closes the dialog box). The **Cancel** button closes the dialog box without executing any of the commands selected.

1. How to Start a New Workbook

When a user initially begins in Microsoft Excel, a new, blank workbook is automatically opened. To begin working, just start typing. A new workbook can be opened anytime Excel is running, even if the current workbook has not been saved or closed. To create a new workbook at any time, simply click the **New Workbook** icon under the Menu bar, or under the **FILE** menu choose the **NEW** command

Even though several workbooks can be open at a given time, all of them may not be seen if the current workbook has not been minimized. To minimize the current workbook, select the down arrow on the Max-Min button that is located on the far right corner of the Menu bar (not to be mistaken for the Title bar). (Refer to the second line or bar of Figure 1-B.2. Notice there is an additional Max-Min button at the far right of this bar.) Once the current workbook has been Minimized, then that workbook's Title bar with its own Maximize and Close buttons can be seen (see Figure 1-C.1.1). The file name for that particular workbook will be located in the middle of the workbook's Title bar (e.g., Book1). To view other current workbooks that have been opened, diminish the workbook to an icon by selecting the Close button, which is the down arrow button, second from the right on the workbook's Title bar. To restore the workbook for viewing, either double-click on the icon or highlight the icon and select the **RESTORE** command.

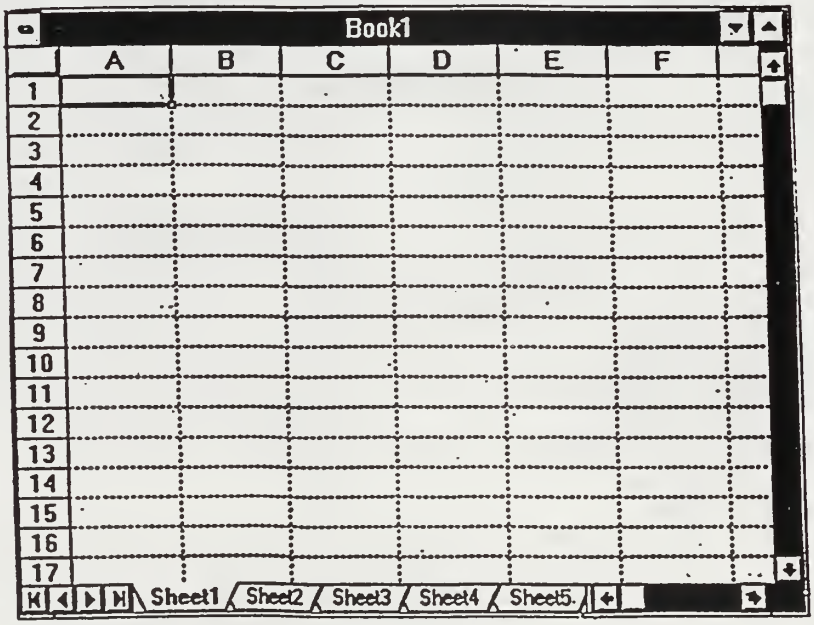


Figure 1-C.1.1: Example of a Blank Workbook

2. How to Save, Open, and Close a Workbook

To save the active workbook, choose the **SAVE** command from the **FILE** menu, or click the **Save** icon. If the workbook has not been saved previously, or if the workbook needs to be saved under a different name than the current one, then select the **SAVE AS** command. The Save As dialog box will appear (see Figure 1-C.2.1) allowing the user to enter a name for the workbook in the File Name text box. Workbook names cannot include any of the following nine characters: ? | * \ / : < > " . The particular drive, directory, or folder in which the workbook is to be saved can also be selected in this dialog box. To ensure the file is saved as a Microsoft Excel workbook, following the file name type ***xls*** after the decimal point (.). Excel automatically adds this suffix to a file name if nothing follows the decimal point. Any suffix added to a Microsoft Excel file name other than the one specified will cause the document to be filed elsewhere. Once the workbook is saved, the new file name entered will show up in the Title bar. A saved workbook's file name will appear in all capital letters with its assigned suffix (e.g., BOOK1.XLS).

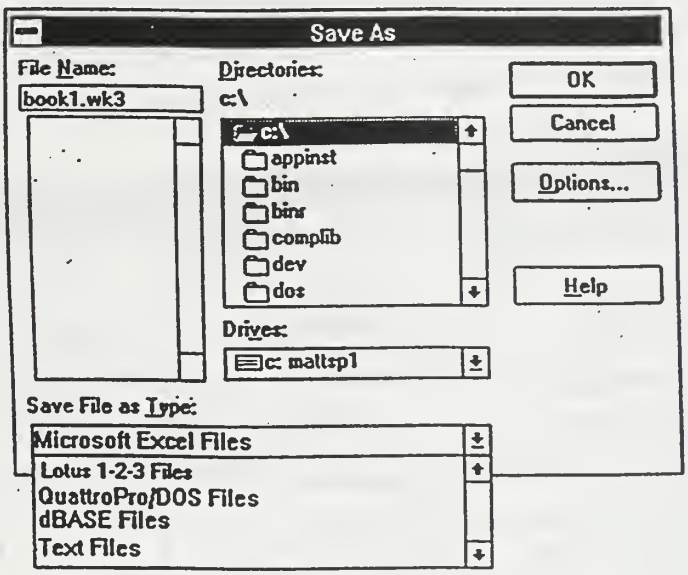


Figure 1-C.2.1: Example of the Save As Dialog Box

Any workbook can be opened by choosing the **OPEN** command from the **FILE** menu or by clicking the **Open** icon in the toolbar. When the Open dialog box appears, select a workbook filename. If a particular workbook cannot be found, due to its filing in a different drive, directory, or folder, then choose the **FIND FILE** button in the Open dialog box to search for it. To remove a workbook from the window simply close it by choosing the **CLOSE** command from the **FILE** menu.

For further information on how to open and save a file from other programs, such as Lotus 1-2-3 and Quattro Pro, see Appendix A: Retrieving and Entering Data from other Files.

3. How to Print a Worksheet

The first step towards printing is to set up the page or the appearance of the printed sheets by selecting the various options in the Page Setup dialog box which is accessed in the **FILE** menu under the **PAGE SETUP** command. Some of the options available to change are: margins, vertical and horizontal alignment on the page, adding or editing headers and footers, and page orientation. Under the **PAGE SETUP** command, by selecting **Sheet tab**, the user can control which parts of the sheet are to be printed, and whether gridlines or row and column headings are to be printed. Also, row or column titles can be specified to appear on multiple pages. If **Page tab** is selected, then such options as enlarging or reducing the printed data, and reducing the data to fit on a specified number of pages, can be selected.

If a sheet is larger than a single page, Microsoft Excel divides it into more pages for printing and inserts automatic page breaks. These page breaks are based on paper size, margin settings, and scaling options from the Page Setup dialog box. Manual page breaks can also be created. However, when a manual page break is set, Microsoft Excel adjusts the automatic page breaks for the rest of the sheet. Dashed lines in a worksheet are used to indicate page breaks. To display page breaks on the screen, choose the **OPTIONS** command under the **TOOLS** menu, and then select the **Automatic Page Breaks** check box (see Figure 1-C.3.1).

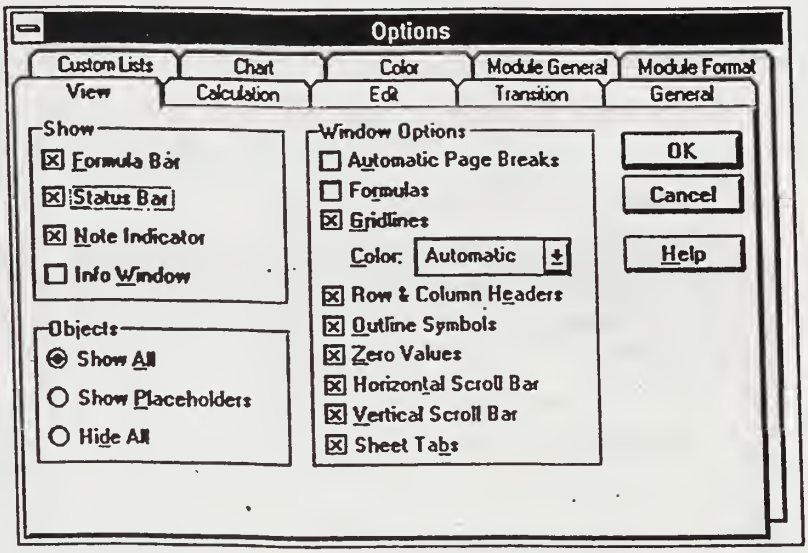


Figure 1-C.3.1: Example of the Options Dialog Box

To select a manual page break creating both a vertical and horizontal break, select the cell where the new page is to begin, and then choose the **PAGE BREAK** command under the **INSERT** menu. However, if only a vertical page break is desired, select the column where the left edge of the new page is to start before choosing the two commands as before. Likewise, if only a horizontal page break is desired, select the row where the top edge of the new page is to start followed by the two commands.

Microsoft Excel numbers and prints pages in one of the following ways:

- **Down, then Across:** Numbering and printing proceed from the first page to the pages below, and then move to the right and continue printing down the sheet. (See Figure 1-C.3.2)
- **Across, then Down:** Numbering and printing proceed from the first page to the pages to the right, and then move down and continue printing across the sheet. (See Figure 1-C.3.2)

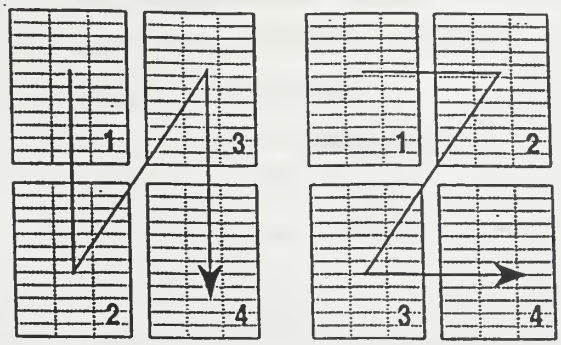


Figure 1-C.3.2: Down, then Across. Across, then Down.

When printing more than one sheet at a time, Microsoft Excel numbers all of the pages in one continuous sequence. To set the page printing order, go to the **FILE** menu, choose the **PAGE SETUP** command, and then select **Sheet tab**. To begin numbering the pages at a specified page number, choose **Page tab** instead of **Sheet tab**.

Before printing a worksheet, it is a good idea to preview the printouts. To preview a worksheet, first open the workbook in which the information is stored. Choose the **PRINT PREVIEW** command from the **FILE** menu, or select the icon with the magnifying glass. The preview command includes many features (such as appropriate margins, page numbers, and titles) that are not shown in the worksheet window.

There are several ways to print the workbook. One can print using the **Print** icon on the standard toolbar, by choosing the **PRINT** command under the **FILE** menu, or by selecting **Print** in the Page Setup dialog box or the Print Preview window. By default, Excel prints the entire, current worksheet if nothing has been highlighted. If only a selected group of cells or rows and columns (e.g., part of a worksheet) needs to be printed, then the user should highlight these cells/rows/columns with the mouse before selecting the Print command.

4. How to Move through a Worksheet

Most worksheets are too large to fit on the screen all at once. By using the horizontal and vertical scroll bars, any part of the worksheet can be brought into view, and any cell activated by a click of the mouse, or one can move into a cell using the arrow keys from the keyboard. There are easier and more efficient ways to move around in a large worksheet. To move immediately to the upper-left corner of the worksheet (usually cell A1), press **Ctrl+Home** from the keyboard. To move to the lower-right cell in the active worksheet, press **Ctrl+End**.

For example, if column F is the last column in the worksheet containing data, and row 10 is the last row containing data, pressing Ctrl+End activates cell F10, regardless of whether that cell contains any data itself. To move to the beginning of the current row, press **Home**. To move to the last non-blank cell in the current row, press **End** followed by **Enter**.

A new workbook opens with 16 sheets named 'Sheet1' through 'Sheet16' located on the sheet tabs at the bottom of the worksheet (see Figure 1-C.4.1). To view quickly or switch between different sheets in a workbook, select a particular sheet to view, and then click on the sheet tab with the mouse to activate that sheet. The scrolling buttons located to the left of the sheet tabs provide an alternative to move through the sheet tabs. However, the scrolling buttons only make different tabs visible. A tab must be highlighted in order to activate a sheet.

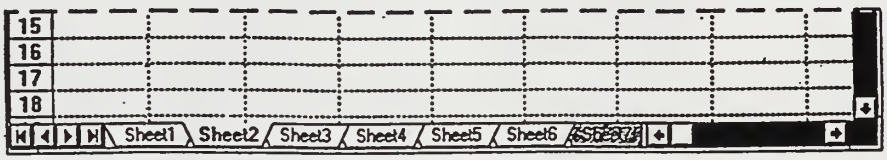


Figure 1-C.4.1: Example of the Sheet Tabs with Scroll Bars

To change the name of any sheet to a name of up to 31 characters, double-click on the sheet's tab. A Rename Sheet dialog box will appear. Once changed, the new name will appear on the sheet tab.

Before data is entered or commands are executed in Excel, a group of cells must be selected. The particular combination of cells selected is called the *range*. It consists of two or more cells that are affected all at once. Ranges can be made up of rows, columns, or rectangles composed of adjacent or nonadjacent cells. One of the most intuitive ways to select cells is to drag across them with the mouse. To do this, first point to the first cell in the desired range. Next hold down the left mouse button and drag to the opposite end or corner of the range. The cells stay highlighted, defining the range. Excel does provide some convenient shortcuts. To select an entire row, simply click on its row number or heading. For example, to highlight Row 3, click on the number *3* just to the left of Row 3; this selects every single cell in that row. To select an entire column, click on its column letter. To select several adjacent columns or rows, then drag across their column letters or row numbers. To select the entire worksheet, click on the **Select All** button which is the blank square located immediately above Row 1 and just to the left of Column A.

D. ENTERING DATA

In general, Excel accepts two types of data: text and numerical values. As a rule, if an entry is not a value, then Excel treats it as text. A value is numeric, including a date, time, currency, percentage, fraction, or scientific notation. When numbers are entered, they align to the right of the cell. If a number is too long to be displayed in

a cell, Excel displays a series of number signs (####) in the cell. If the column is then widened enough to accommodate the width of the number, it is displayed in the cell. (Note: An easy way to widen a column to fit the selection is to double-click the right border of the column heading.) Even if the data does not show in full view, it is nevertheless stored completely in Excel, and Excel uses this complete number when called upon for all calculations (see Figure 1-D.1). Excel stores values with 15 digits of accuracy, called *full precision*. Text can be characters or any combination of numbers and characters. When text is entered, the characters align to the left of the cell. If a text entry is too long for its cell, it will continue over into the adjacent cells if they are empty. Otherwise, the text is cut off at the cell's edge.

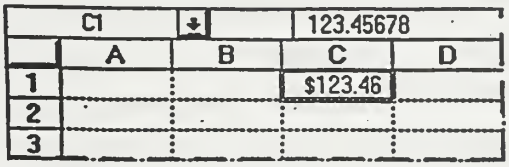


Figure 1-D.1: Example of the Stored Value Displayed in the Formula Bar

Data is typed directly into a cell and does not change unless the cell is edited. Data can be entered into the worksheet whenever the word "*Ready*" appears in the Status bar located at the bottom, right hand corner of the workbook. Entering data actually takes three steps: (1) activate the desired cell or cells, (2) type in the data, and (3) finalize the data entry by pressing **Enter** on the keyboard. Once the desired cell has been made active, type in the data. The Status bar will read "*Enter*" to indicate that data is currently being typed in. Whatever is being entered appears in both the active cell and in the Formula Bar. Three buttons or icons also show up in the Formula bar once the typing begins (see Figure 1-D.2). These buttons (from left to right) are: the **Cancel box** (designated by an 'X') , the **Enter box** (designated by a 'checkmark'), and the **Function Wizard box** (designated by 'fx'). (Note: The Function Wizard is a special tool for building formulas which is addressed in the next section.) Once the typing is done, click the **Enter box** or press the **Enter** key to finalize the data entry. To erase a cell before it has been entered, click **Cancel box** or press the *Escape* (**Esc**) key.

Figure 1-D.2: Example of the Three Icons Located in the Formula Bar

To enter a number as a value, select a cell and type the number. Numbers can include numeric characters (0 through 9) and any of the following characters: + - ( ) , / $ % . E e . Use the following guidelines when entering numbers: commas can be included in numbers, a single period in a numeric entry is treated as a decimal point, plus signs entered before numbers are ignored, and negative numbers must be preceded with a minus sign (or enclosed in parentheses). A number can be quickly formatted as text by preceding it with an apostrophe.

To enter text, select a cell and type the text. A cell can hold up to 255 characters. The characters within the cell can be formatted individually. Instead of allowing long strings of text to overflow into adjacent cells, the text can be displayed on multiple lines within a cell. This format is called *text wrapping*. To activate text wrapping, choose the **CELL** command from the **FORMAT** menu, and select the **Alignment tab** followed by the **Wrap Text** check box.

Excel provides a few neat tricks for speeding up data entry (see Figure 1-D.3). The contents of cells can be copied into other cells by dragging the *fill handle* (the small, black square located in the bottom, right corner of the highlighted cell). Simply select a cell and then point to the fill handle with the mouse. The pointer changes to a thin cross when the mouse pointer is directly over the fill handle. Drag the fill handle to the left, right, up, or down to fill data. When the mouse button is released, the data is filled or copied into the range. A sequence or list can be created by incrementing the value in the active cell by dragging the fill handle through a designated range. For example, a sequence such as 1,2,3 can be extended to include 4,5,6 by using the fill handle.

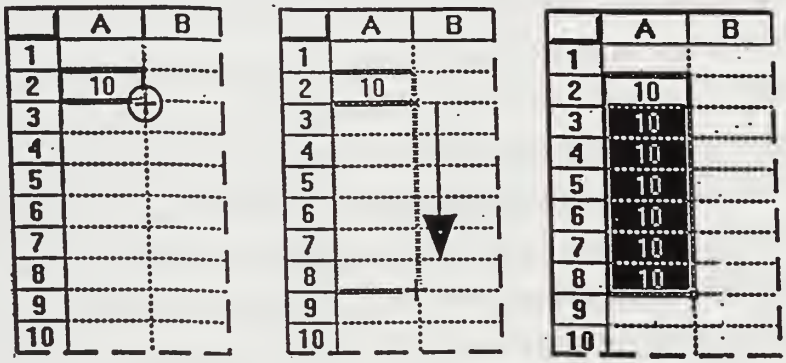


Figure 1-D.3: Example of Dragging to Fill a Range

E. ENTERING FORMULAS

A formula may consist of numbers, text values, functions, or cell addresses combined with arithmetical operators. Using a formula can help analyze the data in a worksheet. An Excel formula always begins with an equals symbol (=). If the equals symbol is not used, Excel will not perform any calculations, and displays instead exactly what was typed. A *constant* is a numeric or text value typed directly into a cell. A simple formula combines constant values with operators (such as plus or minus signs) in a cell to produce a new value from existing values. Think of a formula as one side of an equation whose result is shown in the cell.

As in entering data, a formula is typed directly into a cell and does not change unless it is edited. A formula can be entered into the worksheet whenever the "*Ready*" word appears in the Status bar located at the bottom, right hand corner of the workbook. As before, entering a formula takes three steps: (1) activate the desired cell,

(2) type the formula beginning with an equals symbol (=), and (3) finalize the formula by pressing **Enter** on the keyboard. Once the desired cell has been made active the user can type in the formula; the Status bar reads "*Enter*" to indicate that a formula is currently being entered. Whatever is typed appears in both the active cell and the Formula bar. Likewise, the three buttons or icons show up in the Formula bar once the typing begins (i.e., the **Cancel box** , the **Enter box**, and the **Function Wizard box**). When the typing is complete, click on the **Enter box** or press the **Enter** key to finalize the formula. To erase the cell before it has been entered, click on **Cancel box** or press the *Escape* (**Esc**) key.

A cell containing a formula normally displays the formula's resulting value on the worksheet. When a cell is selected containing a formula, the formula is always displayed in the Formula bar. To set up the worksheet to display formulas instead of resulting values (see Figure 1-E.1), type **Ctrl+`** (left single quotation mark located to the left of the '1' key on the alphanumeric keyboard).

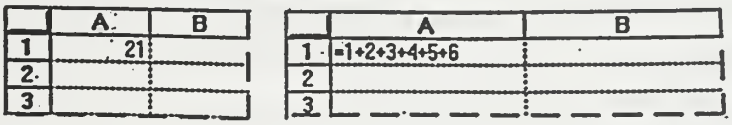


Figure 1-E.1: Example of a Worksheet which Displays Values versus Formulas

Another way to control how formulas are displayed in the worksheet is to use the **OPTIONS** command in the **TOOLS** menu: click **View tab** in the Options dialog box, and then select **Formulas**.

1. Use of Arithmetic and Relational Operators

Operators are used to specify arithmetic operations to be performed on elements of a formula. Arithmetic operators perform basic mathematical operations which combine numeric values and produce numeric results. The following arithmetic operators are used in Excel: + (addition), - (subtraction or negation when placed before a value), / (division), * (multiplication), % (percent which is placed after a value), and ^ (exponentiation). For example, the formula '=20^2*15%' raises 20 to the power of 2 and multiplies the result by 0.15 to produce the result of 60 which is then placed in the active cell.

Relational operators compare two values and produce the logical value *TRUE* or *FALSE*. The following relational operators are used in Excel: = (equal), > (greater than, < (less than), >= (greater than or equal to), <= (less than or equal to), and <> (not equal to). For example, the formula '=A1<25' produces the logical value *TRUE* if cell A1 contains a value less than 25; otherwise, the formula produces the logical value *FALSE*. Finally, there is a text operator ' & ' in Excel to join two or more text values together into a single text string.

If several operators are combined in a single formula, Excel performs the operations in the order shown in Table 1-E.1.1 on the following page:

Operator	Description
-	Negation (as in -10)
%	Percent
^	Exponentiation
* and /	Multiplication and Division
+ and -	Addition and Subtraction
&	Text joining
= < > <= >= <>	Comparison

Table 1-E.1.1: Order of Operations within Excel

If a formula contains operators having the same priority, Excel evaluates them from left to right. To override the order of evaluation, or to ensure a correct calculation, use parentheses to algebraically group expressions in the formula. Excel first calculates the expressions inside parentheses, and then uses those results to calculate the rest of the formula. Do not use parentheses to indicate a negative number within a formula; instead, precede the number with a minus sign.

2. Formula Error Values

When a formula is entered, Excel expects certain types of values for each operator. If something is entered different from what Excel expects to appear, then it tries to convert the value. Excel displays an error value in a cell when it cannot properly calculate the formula for that cell. (Note: If the current formula includes a reference to a cell containing an error value, then the formula also produces an error value.) Error values always begin with a single number sign (#). The following is a list of Microsoft Excel error values together with their meanings:

This Error Value	Means That a Formula
#DIV/0!	is trying to divide by zero.
#N/A	refers to a value that is not available.
#NAME?	uses a name that Excel does not recognize.
#NULL!	specifies an invalid intersection of two areas.
#NUM!	uses a number incorrectly.
#REF!	refers to a cell that is not valid.
#VALUE!	uses an incorrect argument or operand.
#####	produces a result that is too long to fit in the cell. (Also occurs when a constant numeric value is too long. This is not actually an error value, but rather an indicator that the column needs to be wider.)

3. Use of Cell References

Formulas can be made even more powerful using cell references or addresses instead of actual numbers. Cell references are just cell addresses that tell Excel to perform its calculations using the data currently in the designated cell. If the data is changed in a cell referred to by a formula, then the formula automatically changes

the result when new data is entered. This process can be referred to as *automatic recalculation*. With references, data located in different places can be used in one formula, and one piece of data or a single cell's value can be used in several formulas. References identify cells or groups of cells on a worksheet. References tell Excel which cells to read in order to find the values to be used in a formula. Although there are three types of references (relative, absolute, and mixed), the reference most commonly used is the relative one denoted by A1, B2, CC101, etc. An example of a formula using cell references is '=2*(A1+B1)/C1', which means the data in cell A1 is added to the data in cell B1, the result multiplied by two, and finally divided by the data in cell C1.

References can always be entered into formulas, typed in either uppercase or lowercase letters. Excel converts the letters to uppercase when the formula is **Enter**ed. Spaces are not required in formulas, but they may be used to make a formula more readable. Nevertheless, the easiest way to enter references in a formula is to select the cell or range directly on the worksheet. After highlighting the cell in which the formula is to be entered, type an equals symbol and/or an operator (such as +, -, <, or a comma), then click on the cell or drag through the range of cells to be referenced. The selection is surrounded by a dotted line called the *moving border*. The reference to the cell or to the range of cells appears automatically in the formula. Once the cell's reference appears in the formula, **Enter** the formula or continue it by typing another operator.

For example, the formula '=2*(B1+B2+B3)' is to be entered in cell C4. First, highlight cell C4 creating a bold border around the rectangle cell. Now type *=2*(* , but instead of typing B1, go to cell B1 and highlight or activate the cell. Simultaneously, a dotted line or moving border appears around cell B1, and ***B1*** is automatically entered in the formula in cell C4 and in the Formula bar. Now type the plus symbol, +, and continue to add cells B2 and B3 in the same manner as before. Once cell B3 has been included, type in the closing parentheses,), and press **Enter**. The formula is now complete. Although, right now, this procedure appears to be as difficult as directly typing the cell's reference, it will be a necessary technique when several cells are to be included at the same time in a formula.

Also, another means to speed up the process of entering formulas involves the fill handle. Like the data entry process, the contents of cells can be copied into other cells by dragging the fill handle regardless of the cells' contents. That is, if the formulas inputted are similar in structure, except for a change in the cell reference, then the fill handle can be used to copy formulas into other cells. Only formulas that involve incrementing cell references can use this procedure. For example, **Enter** the formula, '=2*A1', in cell B1. If the user wants to multiply the first ten (10) cells in column A by two (2) and have the results appear in column B, then the fill handle can be used to simplify this work. Now that cell B1 has the resulting formula, highlight cell B1, and then point to the fill handle with the mouse. Drag the fill handle down until the end of cell B10. Release the mouse button, and the formula is copied, changed, and automatically updated sequentially into the range. Cell B10 now reflects the formula '=2*A10'.

F. EDITING A WORKSHEET

There are many reasons a user may want to modify the contents of a worksheet: to correct a misspelling, to revise data, to fix a formula entered incorrectly, and so on. Editing cell contents is the same whether the cell contains text, a number, or a formula. There are different ways to edit a cell. The most basic way to edit a cell is to highlight it (or all cells to be edited), and then choose commands from the **EDIT** menu. There are other ways to edit which include the toolbar icons or the shortcut menu (see Figure 1-F.1). The shortcut menu is obtained by first highlighting the cell or cells to be edited, and then clicking the right button on the mouse. A shortened version of the Edit menu appears on the screen with only key commands (e.g., Cut, Copy, Paste, etc.). Select from the available choices by either pointing and clicking with the left mouse button, or by scrolling with the right mouse button. When data has been edited, any formulas relying on that data are recalculated automatically.

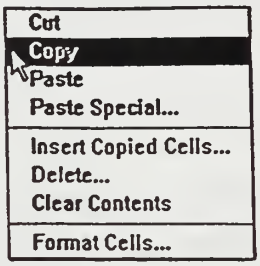


Figure 1-F.1: Example of the Shortcut Edit Menu

1. Modifying the Contents of a Cell

To entirely replace the contents in a particular cell, simply activate the cell to be edited by highlighting it with the mouse (or by using the arrow keys to move onto that cell), and then type in the new data or entry, followed by pressing the **Enter** key. This overwrites the previous contents of the cell. The new data is typed directly over the existing entry, and Excel immediately displays the new data. As before, whenever data is typed, it appears in both the active cell and the Formula bar.

To edit only part of the information within a cell (that is, to modify rather than replace the contents of a cell), double-click on the cell. A flashing insertion line within that cell informs the user that the cell is ready for editing. The Status bar also displays the word "*Edit*". Instead of double-clicking on a cell to display the flashing insertion line, move to the cell to be edited and press the **F2** key. If the cell is displaying formula results from a hidden formula, then the formula now appears and the value is hidden. Now type and delete characters as needed to modify the cell contents. The arrow keys can be used to reposition the flashing insertion line. The **Backspace** key or the **Delete** key can be used to remove text one character at a time. The **Backspace** key deletes characters to the left of the insertion line, while the **Delete** key deletes characters to the right of the insertion point.

Cell contents can also be edited in the Formula bar. The same techniques used to edit within a cell apply to edit within the Formula bar. Once a cell is made active, point and click the cursor in the Formula bar to make the

Formula bar active. If it is desired to make the edits only in the Formula bar, then choose the **OPTIONS** command from the **TOOLS** menu, select **Edit tab**, and then **clear** the **Edit Directly In Cell** check box.

2. Erasing Data from a Worksheet

There are two different ways to go about erasing data from the worksheets. When cells are deleted, they are removed from the worksheet and the surrounding cells shift to fill in the space. When cells are cleared, they are cleared of all content, leaving them empty but remaining in the worksheet. Deleting and clearing cells produces different results for formulas that reference those cells. If a cell's contents are cleared, its value is zero. Thus, a formula referring to that cell receives a value of zero from that cell. If a cell is deleted, it no longer exists; so a formula referring to the deleted cell will not be able to find it and the #REF! Error value is returned.

When data are deleted, any data below or to the right shifts up and/or to the left automatically. When data are shifted, these cells receive new addresses, i.e. if the entire column B is deleted, then the data in column C is moved to column B with a new address beginning with the letter "B". This is important for formulas. Microsoft Excel automatically adjusts formulas that refer to the shifted data so that the formulas continue to produce the correct results.

To delete or clear cells, choose the **DELETE** command or **CLEAR** command from the **EDIT** menu, and then choose from the selections that appear for each command. Here is a confusing quirk in Excel's terminology: the **Delete** key clears cell contents; it does not delete cells from the worksheet. In other words, pressing the **Delete** key does not accomplish the same thing as issuing the **DELETE** command from the **EDIT** menu.

II. MODELING WITH DISCRETE DYNAMICAL SYSTEMS

In Chapter I, the most essential features were discussed to enable the user to work through the mathematical modeling process using Microsoft Excel. This chapter expands on the current working knowledge already presented on Excel to help the user with more complicated and unfamiliar techniques. Although many of the modeling examples in this text can be done without the material that follows, the suggestions presented here will enhance as well as simplify the techniques needed to complete proper models encountered during the mathematical modeling process.

In mathematical modeling, the student is often interested in building models to explain behavior or make predictions. The focus of this chapter is modeling change. By collecting data over a period of time and plotting that data, patterns can be discerned capable of being modeled in a way to capture the trend of the change. This method of modeling a behavior takes place over discrete time periods, and gives rise to a discrete dynamical system.

A. AN EXAMPLE OF FORMULATING DATA

The following example illustrates some of the highlights and features discussed in this chapter.

Example 2.1: The Arms Race

Scenario: Almost all modern wars are preceded by unstable arms races. Strong evidence suggests that an unstable arms race between great powers, characterized by a sharp acceleration in military capability, is an early warning indicator of war. It has been established that rapid competitive military growth is strongly associated with the propensity for predicting war. Thus, by studying the arms race, we have the potential for predicting war (and, possibly, preventing it).

Suppose two countries, Country X and Country Y, are engaged in a nuclear arms race. Assume each country follows a deterrent strategy that requires it to have a given number of weapons (say, missiles) to deter the enemy (inflict unacceptable damage) even if the enemy has no weapons (missiles). Under this strategy, as the enemy adds weapons, the friendly force increases its nuclear arms inventory by some percentage of the number of attacking weapons, which depends on how effective the friendly force perceives the enemy's weapons to be.

Suppose Country Y feels it needs 120 weapons to deter the enemy. Further, for every two weapons possessed by Country X, Country Y feels it needs to add one additional weapon (to ensure 120 weapons remain after a preemptive first strike by Country X). Thus the number of weapons needed by Country Y (y weapons) as a function of the number of weapons it thinks

Country X has (x weapons) is:

$$y = 120 + (1/2)x$$

Now suppose Country X is following a similar strategy, feeling it needs 60 weapons even if Country Y has no weapons. Further, for every three weapons it thinks Country Y possesses, Country X feels it must add one weapon. Thus the number x of weapons needed by Country X as a function of the number y of weapons it thinks Country Y has is:

$$x = 60 + (1/3)y$$

To model the dynamics of the nuclear arms race, suppose initially that Countries Y and X do not think the other side has nuclear arms. Then they build 120 weapons and 60 weapons, respectively. Now assume each has perfect intelligence; that is, each knows the other has built weapons. The nuclear arms race would proceed dynamically; that is, in successive stages. At each stage a country adjusts its inventory based on the strength of the enemy during the *previous* stage. Let n represent the stage of the nuclear arms race and compute the growth of the nuclear arms race under the assumptions stated above.

Using Excel: Open a blank workbook. (Remember, Excel automatically opens a new, blank document when it is started.) The initial number of weapons built by Country Y will be placed in cell B2. (Cell B2 is used in case a heading in cell B1 is desired.) The initial number of weapons is also given its own cell, so the difference equation will be more flexible. Therefore, highlight or select cell B2 with the mouse (or by using the arrow keys) and type **120** (the initial number of weapons built by Country Y). Then press **Enter** or select the **Enter box** (designated by a 'checkmark') located in the Formula bar. The number 120 now appears in cell B2 (see Figure 2.1.1).

	A	B	C	D
1				
2		120		
3				
4				

Figure 2.1.1: The Appearance of the Number 120 in Cell B2

Next, move to or select cell B3. This is where the formula for the dynamical system model or the equation for y will be entered. Once cell B3 has been highlighted, type the following, and then press **Enter**:

=120 + (1/2)*C2

This is the formula for the y equation in Microsoft Excel. The number 120 should appear in cell B3 once the formula has been entered. Since there is no information (e.g., the cell is blank) in cell C2, as referenced in the formula, there is no change in y and the number of weapons for Country Y. (Cell C2

is currently blank, and therefore has 0 (zero) for its value.) This will change once data is placed in cell C2 for Country X.

Now, to enter the data for Country X, the initial number of weapons built by Country X will be placed in cell C2. (Cell C2 is used in case a heading in cell C1 is desired.) The initial number of weapons is also given its own cell, so the difference equation will be more flexible. Therefore, highlight or select cell C2 with the mouse (or by using the arrow keys) and type **60** (the initial number of weapons built by Country X). Then press **Enter** or select the **Enter box** (designated by a 'checkmark') located in the Formula bar. The number 60 now appears in cell C2. Also, notice the data in cell B3 has been automatically updated to reflect the change in cell C2. The number 150 now appears in cell B3. Move back to cell B3 and notice how the numerical answer appears in the cell while the formula appears above in the Formula bar. (see Figure 2.1.2).

	A	B	C	D
1				
2		120	60	
3		150		
4				

Figure 2.1.2: Worksheet Appearance after the Number 60 is Entered in Cell C3

Next, move to or select cell C3. This is where the formula for the dynamical system model or the equation for x will be entered. Once cell C3 has been highlighted, type the following, and then press **Enter**:

=60 + (1/3)*B2

This is the formula for the x equation in Microsoft Excel. The number 100 should appear in cell C3 once the formula has been entered. Move back to cell C3 and notice how the numerical answer appears in the cell while the formula appears above in the Formula bar.

In order to keep track of the number of stages, move to cell A2 and enter the number **0** (zero). The number 0 and the numbers 120 and 60 appear aligned (see Figure 2.1.3). This coincides with the original or initial number of weapons built by Countries Y and X, respectively, before inventories start increasing.

	A	B	C	D
1				
2	0	120	60	
3		150	100	
4				

Figure 2.1.3: Shows the Alignment of the Numbers 0, 120, and 60

Now enter the number **1** (one) in cell A3 to denote the first stage of the nuclear arms race. To keep

track of the first ten stages, highlight both cell A2 and A3 at the same time, by first placing the cursor (or plus sign) on cell A2 and dragging the cursor to cell A3. A sequence will be created by incrementing the column values. Once both cells are highlighted, point to the fill handle at the bottom right corner of the boldface box with the mouse (the pointer changes to a thin cross when the mouse pointer is directly over the fill handle) and drag the fill handle down Column A to the end of Row 12 to fill the data. A faint box should be seen during this process. Release the mouse button and the data 0,1,2,...,10 is filled or copied into Column A, Rows 2 through 12 (see Figure 2.1.4).

	A	B	C	D
1				
2	0	120	60	
3	1	150	100	
4	2			
5	3			
6	4			
7	5			
8	6			
9	7			
10	8			
11	9			
12	10			
13				

Figure 2.1.4: Data 0,1,...,10 is Filled into Column A, Rows 2 through 12

To create the sequence in which the number of weapons for Country Y is shown after intelligence is gathered on Country X and more weapons are built, go to cell B3. Just like before, the automatic entry will be used. But instead of copying data, Excel now copies formulas. This is where the power of Microsoft Excel comes into play and reduces the workload. Point to the fill handle on cell B3 and drag the fill handle down Column B to the end of Row 12. Once the mouse button is released, the cells are filled with the number of weapons after each stage (see Figure 2.1.5).

	A	B	C	D
1				
2	0	120	60	
3	1	150	100	
4	2	120		
5	3	120		
6	4	120		
7	5	120		
8	6	120		
9	7	120		
10	8	120		
11	9	120		
12	10	120		
13				

Figure 2.1.5: Column B Filled with the Number of Weapons after each 10 Stages
(Note: Data for Country X in Column C has not been updated.)

(Remember, the data for Country X in Column C has not been updated, so beginning with cell B5, the formula is reading that Country X has no weapons, as reflected with blank cells, cell C4 through C12.) This will change once data is placed in cells C4 through C12 for Country X. Notice, for example, that cell B4 shows the number of weapons built by Country Y after 2 stages (e.g., Country Y has entered the second stage) while the formula is reflected in the Formula bar.

Now, to create the sequence in which the number of weapons for Country X is shown after intelligence is gathered on Country Y and more weapons are built, go to cell C3. Just like before, the automatic entry will be used. Again, point to the fill handle on cell C3 and drag the fill handle down Column C to the end of Row 12. Once the mouse button is released, the cells are filled with the number of weapons after each stage. Notice that cells C5 through C12 have also been updated automatically to reflect the new data in Column C (see Figure 2.1.6).

	A	B	C	D
1				
2	0	120	60	
3	1	150	100	
4	2	170	110	
5	3	175	116.6667	
6	4	178.3333	118.3333	
7	5	179.1667	119.4444	
8	6	179.7222	119.7222	
9	7	179.8611	119.9074	
10	8	179.9537	119.9537	
11	9	179.9769	119.9846	
12	10	179.9923	119.9923	
13				

Figure 2.1.6: Final Data Results for the Arms Race after 10 Stages

Notice that the growth in the nuclear arms race appears to be diminishing. Numerically, it appears the model can predict an equilibrium value. At this point, the number of stages can be increased, or the initial values for the number of weapons each country first built can be changed in cells B2 and C2. Then the model can be used to test all values for weapons which might be built in a nuclear arms race and to predict an equilibrium value, if it exists. This is just one method of predicting an equilibrium value, by strictly looking at the data. This is a perfect example of how Microsoft Excel decreases the workload and makes changing and testing the model so easy.

B. CREATING CHARTS FROM WORKSHEET DATA

In the mathematical modeling process, it is often desirable to graph a model in order to gain a qualitative feel for interpreting it, or perhaps to use the graph for making predictions. In another situation, the user may wish to test a proposed model against some observed data. Columns and rows of numbers are usually not the best way to

look at the data. It is often helpful to plot the observed data and overlay this scatter plot on a graph of the model. Charts are visual representations of such worksheet data. Microsoft Excel includes several chart types, such as bar, line, and pie charts. A chart is either embedded directly in a worksheet, or it appears on its own chart sheet in the workbook. Excel automatically updates a chart if the worksheet data upon which it is based changes. This section gives the commands necessary to accomplish these tasks with Excel.

A *chart* is a graphical representation of worksheet data. Values from worksheet cells, or *data points*, are displayed as bars, lines, columns, pie slices, or other shapes in the chart (see Figure 2-B.1). However, only scatterplots and line graphs are used for the models.

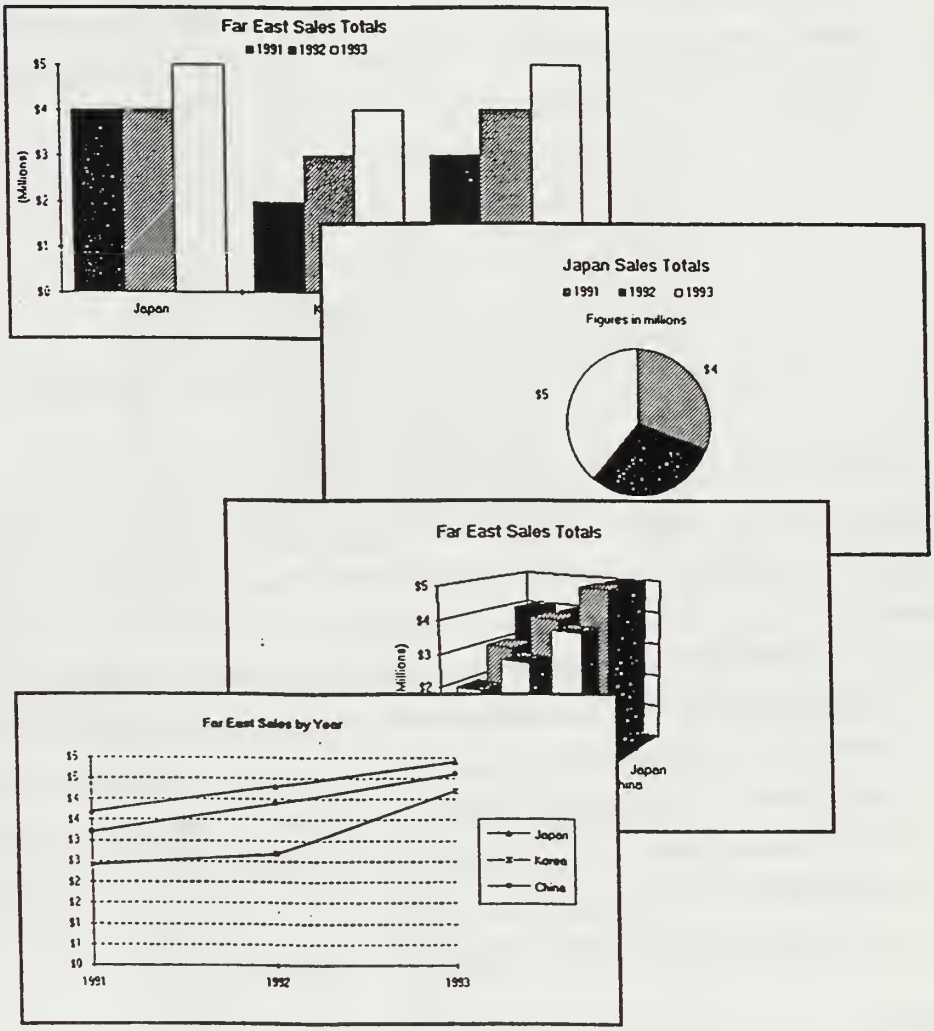


Figure 2-B.1: Example of Several Chart Types

Data points are grouped into *data series*, which are distinguished by different colors or patterns. Showing the data in a scatterplot can make it clearer, more interesting, and easier for the user to discuss trends. Charts also help the user evaluate the data and make comparisons between data values. The *ChartWizard* is a series of

dialog boxes that simplifies creating a chart. It guides the user through the process step by step: verifying the data selection, selecting a chart type, and allowing items to be incorporated such as titles and a legend. A sample of the created chart is displayed on the screen so that changes can be made before exiting the ChartWizard.

1. How to Create a Chart on a Chart Sheet

A *chart sheet* is a special worksheet that contains a chart only. It is created as a separate sheet in a workbook when the user wants to display a chart apart from its associated data (see Figure 2-B.1.1). The chart sheet is automatically linked to the worksheet containing the data. When the data is changed on the worksheet, the chart is updated to reflect these changes. When a chart sheet is created in a workbook, the chart sheet is saved along with the other sheets when the workbook is saved. The chart data is still linked to the source data and is automatically updated when the worksheet data changes. The chart sheet can also be printed independently of the other sheets in the workbook.

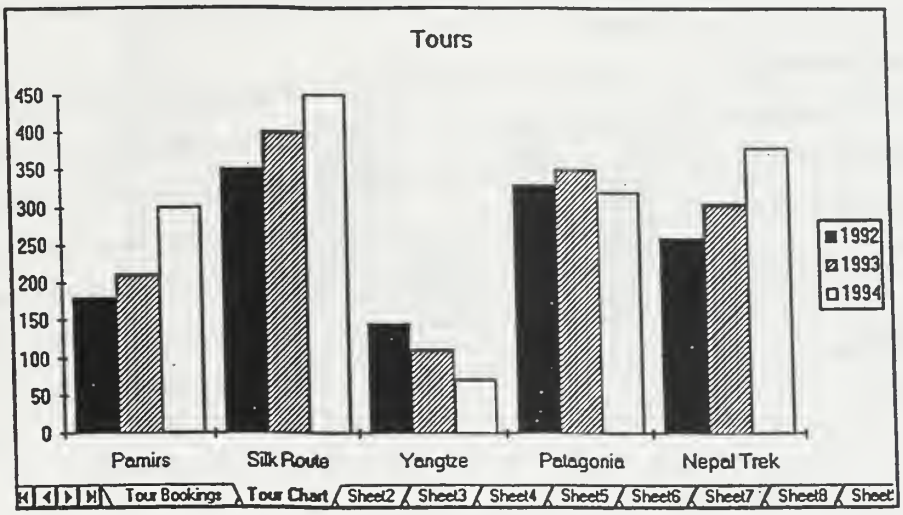


Figure 2-B.1.1: Example of a Chart Sheet

To create a chart sheet, select the worksheet data to be displayed in the chart by highlighting the data. Choose the **CHART** command from the **INSERT** menu, and then choose **As New Sheet** from the submenu that appears. Excel automatically opens the ChartWizard, asking the user to confirm the selected cells. Click on the **Next** button to proceed. Excel lists fourteen (14) available chart types to choose from in picture form. Choose the chart type that presents that particular data most clearly and effectively, and then click on the **Next** button. Each chart type has at least one *subtype*, or variation, which can show the data somewhat differently (see Figure 2-B.1.2). Select one and click on the **Next** button. Review the sample chart in the ChartWizard dialog box, checking whether Excel has interpreted the data correctly. The user may have to tell Excel to take the data series from rows in the worksheet rather than from columns. When complete, click on the **Next** button. Finally, fill in

the information for a legend, a chart title in the Chart Title text box, and the titles for the x-axis and y-axis in the Axis Titles text box, and click on the **Finish** button.

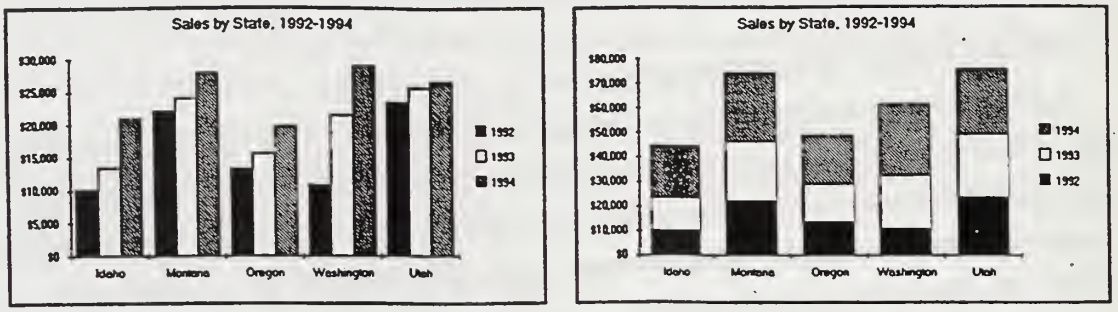


Figure 2-B.1.2: Example of a Chart Type with its Subtype

(Note: A shortcut to creating a chart sheet is to press the **F11** key after selecting the worksheet data. A standard chart is automatically chosen and created.)

Excel adds the new chart sheet to the active workbook, to the left of the worksheet containing the associated data. The chart sheets created in a workbook are named Chart1, Chart2, and so on by default. A chart sheet can be renamed the same way a worksheet: double-click on the sheet tab, type a new name in the Rename Sheet dialog box, and click on the **OK** button. To go back to the sheet containing the data, simply click on its worksheet tab. For example, if the data is stored in Sheet1, then click on the Sheet1 tab after the chart has been added to the workbook. While in the ChartWizard, the **Back** button can be selected to backtrack and make different selections. To exit the chart creation process at any stage, click on the **Cancel** button or press the *Escape* (**Esc**) key.

2. How to Create a Chart on a Worksheet (Embedded)

The user may want to create a chart on an existing worksheet in order to see the chart and the underlying data at the same time. These kinds of charts are sometimes called *embedded charts*. An embedded chart is created as an object on a worksheet to display a chart beside its associated data (see Figure 2-B.2.2). Creating a chart within the current worksheet is not all that different from creating a chart on a sheet of its own. The embedded chart is always available when the worksheet is active, and the chart is printed with the data when the worksheet is printed. The chart data is still linked to the source data and is automatically updated when the worksheet data changes.

As before, select the worksheet data to be displayed in the chart. Click on the **ChartWizard** icon located on the toolbar (see Figure 2-B.2.1), or choose the **CHART** command from the **INSERT** menu and select **On This Sheet** from the submenu that appears. The mouse pointer changes to a cross hair with a tiny chart symbol attached to it. Now, either click anywhere on the worksheet to have the chart placed automatically and at a

Figure 2-B.2.1: The ChartWizard Icon Located on the Toolbar

standard size, or place the cross hair at the corner of a blank cell and drag across the area that will contain the chart. While dragging, a rectangle will appear to indicate the size of the chart. Release the mouse button when the rectangle is the right size. Now follow the instructions in the ChartWizard as before, clicking on the **Next** button to proceed or the **Back** button to return to the previous dialog box. After clicking on the **Finish** button, the chart will appear on the worksheet.

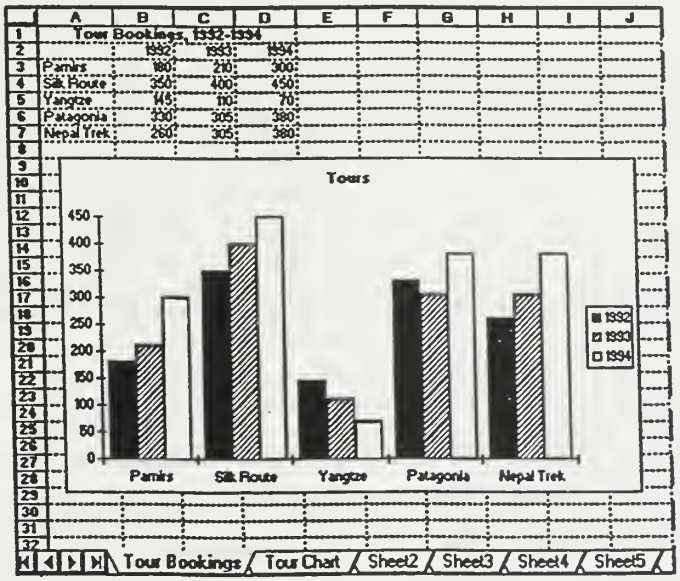


Figure 2-B.2.2: Example of a Chart on a Worksheet (Embedded Chart)

a. An Example of Creating a Chart

The following example illustrates some of the highlights and features currently under discussion. Refer to Chapter II, Section A to review how the data was obtained. This example gives the step-by-step process for creating a chart on a worksheet (embedded chart).

Example 2.2(a): The Arms Race Revisited

Scenario: Recall, Country X and Country Y are engaged in a nuclear arms race. Each country is following a deterrent strategy that requires it to have a given number of weapons to deter the enemy (inflict unacceptable damage) even if the enemy has no weapons. Under this strategy, as the enemy adds weapons, the friendly force increases its nuclear arms inventory by some percentage of the number of attacking weapons, which depends on how effective the friendly

force perceives the enemy's weapons to be. Country Y feels it needs 120 weapons, initially, to deter the enemy. Country X feels it needs 60 weapons, initially, even if Country Y has no weapons. They build 120 weapons and 60 weapons, respectively. The nuclear arms race now proceeds in successive stages. At each stage, a country adjusts its inventory based on the strength of the enemy during the *previous* stage. The growth of the nuclear arms race is computed for 10 stages (see Figure 2.2.1). This trend is reflected in a graph.

Stage n	Country Y	Country X
0	120.00	60.00
1	150.00	100.00
2	170.00	110.00
3	175.00	116.67
4	178.33	118.33
5	179.17	119.44
6	179.72	119.72
7	179.86	119.91
8	179.95	119.95
9	179.98	119.98
10	179.99	119.99

Figure 2.2.1: Data Results for Country Y and Country X after 10 Stages

Using Excel: With the workbook open to the data above, select or highlight Columns A, B, and C, including only Rows 2 through 12. (This is the data to be displayed in the chart.) Now, click on the **ChartWizard** icon located on the toolbar, or choose the **CHART** command from the **INSERT** menu and select **On This Sheet** from the submenu that appears. (The mouse pointer changes to a cross hair with a tiny chart symbol attached to it.) Place the cross hair in the middle of cell D1 and click the right mouse button. (This is where the upper right corner of the chart will be placed.) Immediately, the ChartWizard dialog box opens in a 5 step process.

Step 1 asks to verify the range. The following should appear in the range box: **=$A$2:$C$12** (given the data was input as above). (Note: The dollar sign (**$**) before the column letter and the row number indicates that Excel is using an absolute cell reference. This particular reference tells Excel how to find a cell based on the exact location of that cell in the worksheet.) If changes need to be made to the range, make them now. If the range is correct, select the **Next** button. The Step 2 dialog box now appears and asks for a chart type. Select the chart type **Line**, and press the **Next** button.

Step 3 asks to select a format for the Line chart. Select format **10**, which represents a smooth curve, then press the **Next** button. Step 4 asks several questions, but the only correction that is necessary is to change 0 (zero) to **1** (one) in: *Use First __ Column(s) for Category(X) Axis Labels*, then press the **Next** button. By making this correction, Column A (which is the first

column of data) is now designated as the range for the x-axis, and therefore is not plotted. Before, Column A (which is the number of Stages in the nuclear arms race) was highlighted as part of the range, and therefore was plotted on the graph along with the other data. This is an easy way of designating the values for the x-axis without much work. Finally, Step 5 asks for a Legend, Chart Title, and Axis Titles. Once these labels are typed, select the **Finish** button, which automatically closes the ChartWizard dialog box and enters a chart on the worksheet next to the data (see Figure 2.2.2).

	A	B	C	D	E	F	G	H	I
1	Stage n	Country Y	Country X						
2	0	120.00	60.00						
3	1	150.00	100.00						
4	2	170.00	110.00						
5	3	175.00	116.67						
6	4	178.33	118.33						
7	5	179.17	119.44						
8	6	179.72	119.72						
9	7	179.86	119.91						
10	8	179.95	119.95						
11	9	179.98	119.98						
12	10	179.99	119.99						
13									

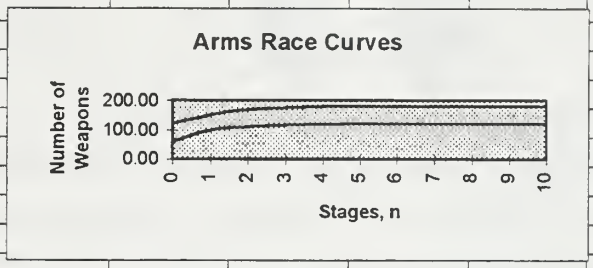


Figure 2.2.2: Rough Arms Race Chart Next to the Data

(Remember, if any mistakes or changes are made during the ChartWizard's 5 Step process, press **Cancel** to exit and close the ChartWizard dialog box.) The chart may appear rough at first (see Figure 2.2.2), but editing can be made directly on the chart itself as explained in the following section.

3. How to Modify a Chart

To make changes to a chart, the user must first activate it. To modify a chart on a chart sheet, activate the sheet containing the chart to be changed by selecting the appropriate sheet/chart tab. Now choose **CHART TYPE** from the **FORMAT** menu, and a Chart Type dialog box will appear with several options. Once the appropriate selections have been made, choose the **OK** button to activate the new choices. If the Chart toolbar is displayed, then changing the chart type can be done quick and easy. Often the Chart toolbar is displayed when a chart sheet is created. If the Chart toolbar is not visible, it can be redisplayed by choosing **TOOLBARS** from the **VIEW** menu, selecting the **Chart check box**, and clicking on the **OK** button.

To modify a chart that is embedded within a worksheet, select the embedded chart by clicking on it once. This will enclose the embedded chart with a set of square handles. Double-clicking on an embedded chart activates it, which gives it a set of square handles and changes the chart border from a thin line to a thicker gray border (see Figure 2-B.3.1). At this point, either choose a new chart type from the **Chart Type** icon on the

Chart toolbar, or choose **CHART TYPE** from the **FORMAT** menu, make a selection from the Chart Type dialog box, and click the **OK** button. Only changes can be made to an embedded chart when the embedded chart is active.

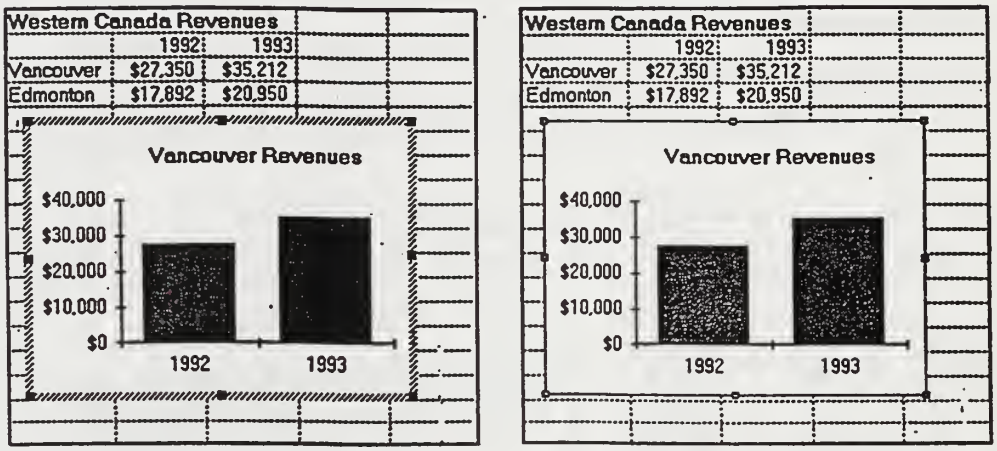


Figure 2-B.3.1: An Active Embedded Chart (*left*) vs. A Selected Embedded Chart (*right*)

To alter a chart on a chart sheet, activate the sheet containing the chart. At this point, to add items to a chart, such as a new chart title or a legend, use the commands on the **INSERT** menu and make the appropriate selections. Or, to change a chart or axis title directly, simply select that item on the chart, which encloses it with a set of square handles. Type the new title and press **Enter**. To alter a chart embedded within a worksheet, double-click on the chart to activate it, and then add or edit titles as mentioned above.

The following information pertains to either type of chart. To delete an item, such as a legend or a title, simply click on that item and then press the **Delete** key. To remove the set of square handles, simply press the *Escape* (**Esc**) key, or click anywhere outside the chart. If a chart is active, that is enclosed within a heavy gray border, click anywhere outside the chart to remove the border but not the square handles. To undo the selection of a chart or a chart item, press the *Escape* (**Esc**) key.

To move or resize a chart, select the chart, ensuring it is enclosed in a set of selection handles. To move the chart, place the mouse pointer anywhere over the chart. The pointer will change into an arrow. Drag the chart to a new location. While dragging, a dashed rectangle indicates where the chart will be placed if the mouse button is released. Release the mouse button when the chart has reached its new location. To resize a chart, again making sure that the chart is only selected, place the mouse pointer over one of the side or corner selection handles. The pointer changes into a double-headed arrow. Drag to shrink or enlarge the chart. Again, a dashed rectangle shows the new size of the chart. Release the mouse button when the chart has the right size.

Example 2.2(b): The Arms Race Revisited (continued)

Using Excel: To edit the chart in order to improve its appearance, click on the chart, enclosing the embedded chart with a set of black, square handles. (Note: Most edit commands will be found under the **FORMAT** menu.) First, to resize the chart, place the mouse pointer over one of the side or corner selection handles (the pointer changes into a double-headed arrow) and drag to enlarge the chart. Excel automatically has a page break every nine (9) columns, so be sure not to cross the dotted line between Columns H and I, which is the page break indicator. Enlarge the chart down to the bottom of Row 13. While the border still has the square, selection handles, choose **OBJECT** from the **FORMAT** menu in order to modify the border to the chart diagram. In the Format Object dialog box, choose the **Patterns tab**, and then select **None** under the Border menu. Press the **OK** button to exit, which implements the choices selected. In this case, the border is erased.

Now, to edit inside the chart, double-click on the embedded chart to activate it, which gives it a set of black, square handles and changes the chart border to a thick, dashed gray or blue border. Now that the chart is activated, any components within the chart can be selected and edited. Select the gray, plot area on the chart and choose **SELECTED PLOT AREA** under the **FORMAT** menu. A Format Plot Area dialog box will appear. Choose **None** under the Borders menu and **None** under the Area menu, then press the **OK** button. These commands erase the gray background in the plot area, as well as, the border on the graph itself.

Next select the Chart Title with the selection handles and choose **SELECTED CHART TITLE** under the **FORMAT** menu. The Format Chart Title dialog box appears, so choose the **Font tab**. Under the Font tab, change the Font Style from Bold to **Regular** and press the **OK** button. Make the same editing changes for the axis titles or labels.

Finally, the x- and y-axes need modifying. First, select the y-axis with the selection handles, then choose **SELECTED AXIS** under the **FORMAT** menu. The Format Axis dialog box appears, so choose the **Number tab**. Under the Number tab, change the Format Code from 0.00 to **0** (zero). Now, instead of selecting the OK button, choose the **Patterns tab** and change the Tick Mark Type: Major to **Inside**. Now press the **OK** button, which implements all the changes made under the Format Axis dialog box at once. Select the x-axis, and make similar changes. One additional change to the x-axis can be found under the **Alignment tab** in the Format Axis dialog box. Choose the top **Text box** in order to rearrange the direction the numbers are facing along the x-axis. Press the **OK** button, and the chart is complete (see Figure 2.2.3). To remove the border and the selection handles from around the chart, double-click anywhere outside the chart. Now the chart is easier to read (see Figure 2.2.3).

Stage n	Country Y	Country X
0	120.00	60.00
1	150.00	100.00
2	170.00	110.00
3	175.00	116.67
4	178.33	118.33
5	179.17	119.44
6	179.72	119.72
7	179.86	119.91
8	179.95	119.95
9	179.98	119.98
10	179.99	119.99

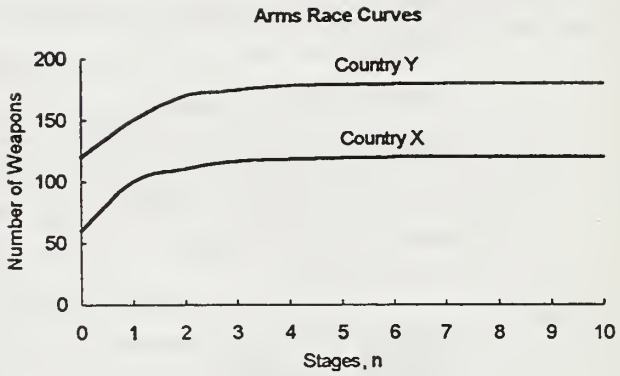


Figure 2.2.3: Modified Arms Race Chart with Data

(Note: To obtain the data curve labels, select any single point on each curve, respectively. Then choose the **SELECTED DATA POINT** under the **FORMAT** menu. When the Format Data Point dialog box appears, mark **Show Label** under the **Data Labels tab**, and after pressing the **OK** button, the data point's coinciding label number appears beside the selected data point. Now, highlight the label, and edit it directly by typing **Country X** or **Country Y** in place of the number. Repeat this procedure for the other curve.)

4. How to Add and Delete Chart Data

Once a chart is created, it is sometimes necessary to update the chart by adding or deleting data series or data points. In some cases, the range of the worksheet data the chart is based on may need to be changed. To add data to a chart sheet, select the worksheet data to be included, and then choose the **COPY** command from the **EDIT** menu. Activate the chart sheet containing the chart and choose the **PASTE** command from the **EDIT** menu. To add data to an embedded chart on a worksheet, select the additional data on the worksheet and use the mouse pointer to drag the selected data over the chart. When the mouse pointer turns into an arrow with a plus sign attached, release the mouse button to "drop" the new data onto the chart. The new data is automatically incorporated into the chart. For either type of chart, new data can be added using the **NEW DATA** command on the **INSERT** menu. To use this command, activate the chart to add the data, choose **NEW DATA** from the **INSERT** menu, type or drag across the range containing the data to be added to the chart, and click on the **OK** button.

To delete a data series or data points, select the data from the worksheet, and then press the **Delete** key. Another method to delete a data series or set of points is to select the data from the worksheet, choose the **CLEAR** command from the **EDIT** menu, and then make the appropriate choice from the submenu. To completely change the worksheet range or display a set of different data, select the embedded chart on the

worksheet or switch to the chart sheet in the workbook, and then click the **ChartWizard** icon. In Step 1 of the ChartWizard dialog box, specify the new range to be plotted in the chart.

Example 2.2(c): The Arms Race Revisited (continued)

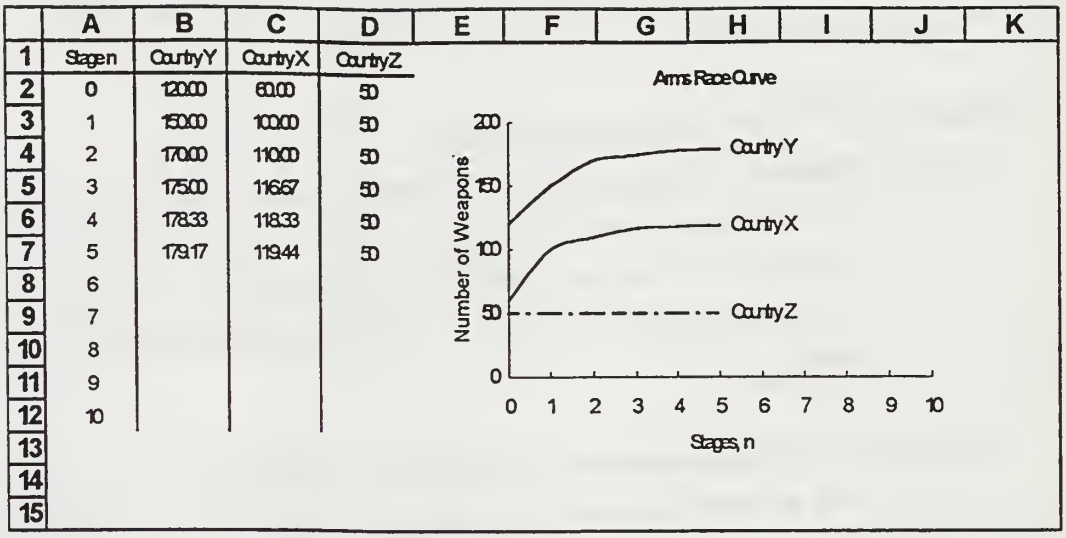
Scenario: (*continued*) A third country, Country Z, has entered the nuclear arms race with Countries X and Y. However, its strategy is to maintain 50 weapons at all times no matter how many weapons the enemy may have. Therefore, Country Z will have 50 weapons during every stage of the nuclear arms race. Now reflect the trend of Countries X, Y, and Z after only 5 stages on a graph.

Using Excel: Add Country Z to Column D, beginning with the number **50** in cell D2. Use the Automatic Fill Handle and drag the data down Column D to the end of Row 7. Once the mouse button is released, the cells are filled with the number 50 after 5 stages.

To add Country Z or the new data to the embedded chart, activate the chart by double-clicking onto the chart. (If the chart is large enough, instead of a border being generated around the chart, the chart is placed inside its own window. Treat the chart in the window the same as a chart with a border.) Once the chart is activated, select **NEW DATA** from the **INSERT** menu, and a New Data dialog box appears. Either enter the range in the space provided, or click on any part of the worksheet to activate the worksheet, then go to cell D2 and highlight it. A dotted rectangle will border the cell, then drag the dotted rectangle to the end of cell D6, and press the **OK** button. This initiates a Paste Special dialog box to appear. Mark **New Series** for "Add Cells as", and mark **Columns** for "Values (Y) in". (If these were additional points to be added to a series already graphed, then **New Point(s)** would be marked for "Add Cells as".) Also, if the range of the x-axis (e.g., Column A or the number of Stages) is included or highlighted with the Country Z data points, then **Categories (X Labels) in First Column** would be marked as well to identify Column A as the x-axis points. Once the **OK** button is pressed, the data is drawn on the graph.

To clear the graph of the data points from Country X and Y for stages 6 through 7, highlight cells B8 through B12 and cells C8 through C12. Then, from the **EDIT** menu, choose **CLEAR** followed by either **ALL** or **CONTENTS** from the submenu, depending on if the formatting for those cells needs to remain in place. The chart is automatically updated and changed, and it should appear similar to Figure 2.2.4.

With the tools outlined in this section, the user is well on his or her way to make charts which effectively present the data.



	A	B	C	D
1	Stage n	CountryY	CountryX	CountryZ
2	0	120.00	60.00	50
3	1	150.00	100.00	50
4	2	170.00	110.00	50
5	3	175.00	116.67	50
6	4	178.33	118.33	50
7	5	179.17	119.44	50
8	6			
9	7			
10	8			
11	9			
12	10			
13				
14				
15				

Figure 2.2.4: Modified Chart with Countries X, Y, and Z

C. MODELING CHANGE WITH DIFFERENCE EQUATIONS

In this section, mathematical models are built to describe the change in an observed discrete behavior. A mathematical model written using Microsoft Excel allows the user to experiment mathematically with different conditions that affect the model, to better understand the behavior under investigation. The following example demonstrates the use of Microsoft Excel to model a behavior in the real world with a difference equation.

Example 2.3: Mortgaging a Home

Scenario: Six years ago you purchased a home by financing $80,000 for 20 years, paying monthly payments of $880.87 with a monthly interest of 1%. You have made 72 payments and wish to know how much you owe on the mortgage, which you are considering paying off with an inheritance you received. (Or you could be considering refinancing the mortgage with several interest rate options, depending on the length of the payback period.) The change in the amount owed each period (e.g., a period, in this case, is a month) increases by the amount of interest and decreases by the amount of the payment:

$$\Delta b_n = b_{n+1} - b_n = 0.01b_n - 880.87$$

Solving for b_{n+1} and incorporating the initial condition gives the dynamical system model:

$$b_{n+1} = b_n + 0.01b_n - 880.87,$$
$$b_0 = 80000,$$

where b_n represents the amount owed after n months.

Using Excel: To create the data, enter the original purchase price of the home b_0 in cell C3. (Cell C3 is used in case a heading in cells C1 and C2 is desired.) The original purchase price is also given its own cell, so that the difference equation will be more flexible. Therefore, highlight or select cell C3 with the mouse, type **80000** (which is b_0 or the original purchase price), and then press **Enter**. The number 80000 now appears in cell C3.

Next, move to or select cell C4. This is where the formula for the dynamical system model or b_{n+1} will be entered. Once cell C4 has been highlighted, type the following, and then press **Enter**:

=C3 + (0.01*C3) - 880.87

This is the formula for b_{n+1} in Microsoft Excel. The number 79919.13 should appear in cell C4 once the formula has been entered. Move back to cell C4 and notice how the numerical answer appears in the cell while the formula appears above in the Formula bar.

In order to keep track of the number of monthly payments, move to cell B3 and enter the number **0** (zero). The number 0 and the number 80000 appear side-by-side (see Figure 2.3.1). This coincides with the original purchase price of $80,000 with zero (0) payments having been made.

	A	B	C	D
1				
2				
3		0	80000	
4		1	79919.13	
5		2	79837.45	
6		3	79754.96	
7		4	79671.64	
8		5	79587.48	
9		6	79502.49	
10		7	79416.64	
11		8	79329.94	
12		9	79242.37	
13		10	79153.92	
14		11	79064.59	
15		12	78974.37	
16				

Figure 2.3.1: Final Data Results for Mortgaging a Home After 12 Payments

Now enter the number **1** (one) in cell B4 to denote the first payment of $880.87. To keep track of the first twelve payments, highlight both cell B3 and B4 at the same time, and drag the fill handle down Column B to the end of Row 15 to fill the data. A faint box should be seen during this process. Release the mouse button and the data **0,1,2,..,12** is filled or copied into Column B, Rows 3 through 15 (see Figure 2.3.1).

To make the sequence in which the amount owed is shown after each payment is made, go to cell C4. Just like before, the automatic entry is used, but to copy the formula instead of data. This is where the power of Microsoft Excel comes into play and reduces the workload. Drag the fill handle

down Column C from cell C4 to the end of Row 15. Once the mouse button is released, the cells are filled with the amount owed after each payment (see Figure 2.3.1).

To make the data more presentable and easier to read, add column headings to correctly identify the data. Therefore, highlight cell B1, type **Months**, and then press **Enter** or select the **Enter box** (designated by a 'checkmark') located in the Formula bar. Next, type **n** in cell B2, **Amount Owed** in cell C1, and **bn** in cell C2. Then press **Enter** (see Figure 2.3.2).

	A	B	C	D
1		Months	Amount Owed	
2		n	bn	
3		0	80000	
4		1	79919.13	
5		2	79837.45	
6		3	79754.96	

Figure 2.3.2: Column Headings Before Editing

To edit the format of the data, highlight Columns B and C, Rows 1 through 15, and then choose **CELLS** under the **FORMAT** menu. A Format Cells dialog box will appear. Under the **Alignment tab**, change the Horizontal alignment to **Center**, change the Font Size from 10 to **8** under the **Font tab**, and then press the **OK** button. Now, highlight Column C, Rows 3 through 15, and then go back into the **CELLS** command under the **FORMAT** menu. Under the **Numbers** tab, change General to **0.00**; then press the **OK** button. Now the data is aligned and centered (see Figure 2.3.3).

	B	C
1	Months	Amount Owed
2	n	bn
3	0	80000.00
4	1	79919.13
5	2	79837.45
6	3	79754.96
7	4	79671.64
8	5	79587.48
9	6	79502.49
10	7	79416.64
11	8	79329.94
12	9	79242.37
13	10	79153.92
14	11	79064.59
15	12	78974.37

Figure 2.3.3: Editing Changes to the Data

With the workbook open to the data below, an embedded chart will be created to display both the data and the chart on the same worksheet. Select or highlight Columns B and C, including only Rows 3 through 15. (This is the data to be displayed in the chart.) Next, click on the **ChartWizard**

icon located on the toolbar, place the cross hair in the middle of cell D1, and click the right mouse button. (This is where the upper right corner of the chart will be positioned.) Immediately, the ChartWizard dialog box opens in a 5 step process.

Step 1 asks to verify the range. The following should appear in the range box: **=$B$3:$C$15** (given the data was input as above). (Note: Again, the dollar sign (**$**) in the range indicates Excel is using an absolute reference for the location of the cells in the worksheet.) If changes need to be made to the range, then this is the time to make those changes. If the range is correct, select the **Next** button. The other four steps are self-explanatory and should include: **Line** chart type, format **10** for smooth curve, changing 0 (zero) to **1** (one) in: *Use First __ Column(s) for Category(X) Axis Labels* (to designate the first column as the range for the x-axis, and therefore not be plotted). Finally, supply the information for a Legend, Chart Title, and Axis Titles. Once these labels are typed, select the **Finish** button, which automatically closes the ChartWizard dialog box and enters a chart on the worksheet next to the data (see Figure 2.3.4). The chart may appear rough at first, but editing can be made directly on the chart itself, if needed.

To edit the chart in order to improve its appearance, double-click on the chart to activate it, enclosing the embedded chart with a set of black, square selection handles and changes the chart border to a thick, dashed gray or blue border. (If the chart is quite large, the chart will be placed in its own window instead of using a border while formatting is changed within the chart.) Now that the chart is activated, all components within the chart can be selected and edited accordingly. Most changes to a chart include: resizing the chart, erasing the gray plot area, changing font style for the titles, and modifying the x- and y-axes. (Note: Most design and appearance changes can be found under the **FORMAT** menu once an item is selected.) The chart is now complete.

Months n	Amount Owed b_n
0	80000.00
1	79919.13
2	79837.45
3	79754.96
4	79671.64
5	79587.48
6	79502.49
7	79416.64
8	79329.94
9	79242.37
10	79153.92
11	79064.59
12	78974.37

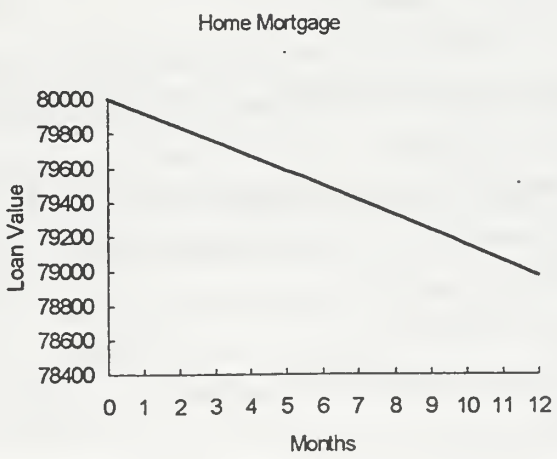


Figure 2.3.4: Modified Home Mortgage Chart with Data

To remove the window and the selection handles from around the chart, double-click anywhere outside the chart. Now the chart is easier to read (see Figure 2.3.4).

Since the home mortgage is for 20 years, to find the amount owed after every payment, the months have to be extended from 12 months to 240 months (e.g., 20 years x 12 months/year = 240 months). The Amount Owed (Column C) also needs to be extended through the same 240 months. Once this data has been extended through 240 months, it can be added to the chart. (To view a copy of this data, see Appendix B: Numerical Data for Mortgaging a Home). To add data to the chart, activate the chart by double-clicking on the chart. Then, choose **NEW DATA** from the **INSERT** menu, type or drag across Columns B and C, Rows 16 through 243, which is the new data to be added, and click on the **OK** button. The chart is now updated to reflect payments throughout the entire 240 months (Figure 2.3.5).

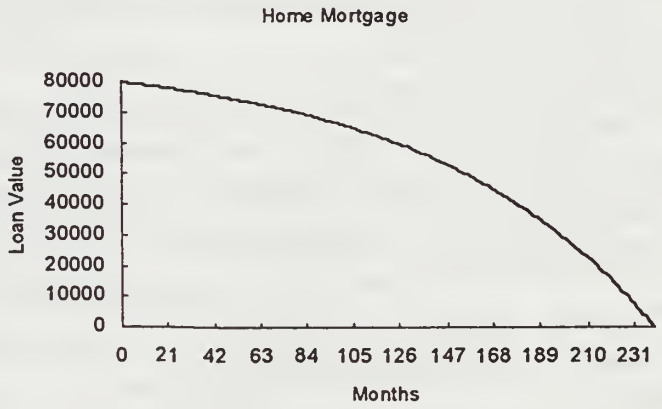


Figure 2.3.5: Example of the Complete Home Mortgage Chart

The updated chart can be modified to improve its presentation. First, select the x-axis, and then press **SELECTED AXIS** under the **FORMAT** menu. Under the **Scale tab**, change 1 to **21** in: *Number of Categories between Tick-Mark Labels*, change 1 to **21** in: *Number of Categories between Tick Marks*, and then press the **OK** button. This eliminates the heavy border on the x-axis. Now select the y-axis, and again choose **SELECTED AXIS** under the **FORMAT** menu. Under the **Scale tab**, change *Minimum* from - 10000 to **0** (zero), and then press the **OK** button. This eliminates the y-axis labels below 0 (zero). Finally, the only step left to make the chart more presentable is to extend the right border of the chart outward until there is enough room for the x-axis labels to come on line.

D. APPROXIMATING CHANGE WITH DIFFERENCE EQUATIONS

Describing some change mathematically is not as precise a procedure as in the case of the home mortgage example. Typically, the user plots the change in data measurements, observes a pattern, and then approximates the change pattern in mathematical terms. Modeling change is the art of determining or approximating the observed pattern of change. Few models capture exactly the real world. Generally, a mathematical model simply approximates real-world behavior. That is, some simplification is required to represent a real-world behavior with a mathematical construct. Here is a simple example of a constructed model.

Example 2.4: Growth of a Yeast Culture

Scenario: Consider the data in Figure 2.4.1 collected from an experiment measuring the growth of a yeast culture. The graph suggests that the change in population is proportional to its current size. Symbolically,

$$\Delta p_n = p_{n+1} - p_n = kp_n$$

where p_n represents the size of the biomass after n hours and $k > 0$ is a constant of proportionality.

Time in hours n	Observed yeast biomass p_n	Change in biomass $p_{n+1} - p_n$
0	9.6	
1	18.3	8.7
2	29.0	10.7
3	47.2	18.2
4	71.1	23.9
5	119.1	48.0
6	174.6	55.5
7	257.3	82.7

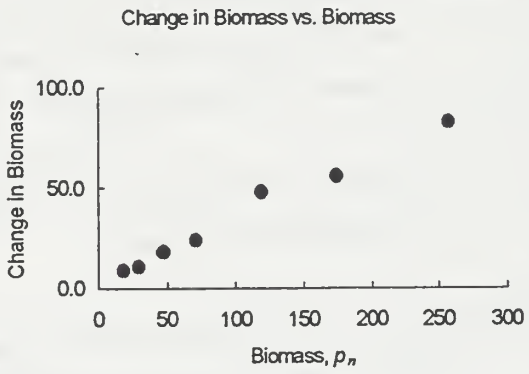


Figure 2.4.1: Growth of a Yeast Culture Versus Time in Hours

Using Excel: To create Figure 2.4.1, go to **Sheet1** and input the data for n and p_n given in the two columns, beginning with cells A5 and B5, respectively. Then create the data for $p_{n+1} - p_n$ for the third column by placing the following Excel formula in cell C6:

=B6−B5

The data should appear similar to Figure 2.4.2, with some editing preferences.

To create a chart on the worksheet, highlight Columns B and C for Rows 6 through 12 (the only data that is needed). Column B (or p_n) will be the values for the x-axis, and Column C (or $p_{n+1} - p_n$) will represent the values for the y-axis. Once in the ChartWizard dialog box, choose chart type

(XY) Scatter, and then select format **1** (one) for the (XY) Scatter chart (which gives individual markers without any connecting curves or gridlines). With several editing options, the chart appears similar to that in Figure 2.4.2 below.

	A	B	C	D
1	Time	Observed	Change	
2	in	yeast	in	
3	hours	biomass	biomass	
4	n	p_n	$p_{n+1} - p_n$	
5	0	9.6		
6	1	18.3	8.7	
7	2	29.0	10.7	
8	3	47.2	18.2	
9	4	71.1	23.9	
10	5	119.1	48.0	
11	6	174.6	55.5	
12	7	257.3	82.7	
13				

Figure 2.4.2: Data for Growth of a Yeast Culture

Scenario: (*continued*) Although the graph of the data does not lie exactly along a straight line passing exactly through the origin, we see it can be *approximated* by such a straight line. The slope is estimated to be about 0.5. Using this estimate $k = 0.5$, the proportionality model is represented by

$$\Delta p_n = p_{n+1} - p_n = 0.5p_n$$

to yield the predictions $p_{n+1} = 1.5p_n$. Note that the model predicts a population that increases forever (e.g., without bound).

Using Excel: To graph this predicted line or data onto the original chart, add a fourth column (Column D) to the table, and type in the following Excel formula starting with cell D6:

=0.5*B6

Column D represents the predicted change in biomass, where $\Delta p_n = 0.5p_n$ (see Figure 2.4.3). Then just add the new data points onto the chart using different shaped markers or a line (see Figure 2.4.4). From the new graph, it is easy to see that the predicted line is just an approximation, but it is not very accurate.

	A	B	C	D	E
1	Time	Observed	Change	Predicted	
2	in	yeast	in	Change in	
3	hours	biomass	biomass	biomass	
4	n	p_n	$p_{n+1} - p_n$	$0.5p_n$	
5	0	9.6			
6	1	18.3	8.7	9.2	
7	2	29.0	10.7	14.5	
8	3	47.2	18.2	23.6	
9	4	71.1	23.9	35.6	
10	5	119.1	48.0	59.6	
11	6	174.6	55.5	87.3	
12	7	257.3	82.7	128.7	
13					

Figure 2.4.3: Experimental Data and Predicted Data for Growth of a Yeast Culture

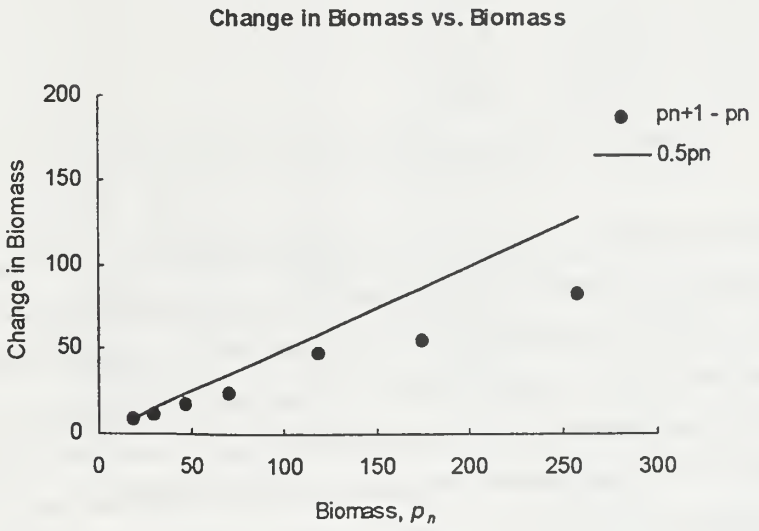


Figure 2.4.4: Plot of the Yeast Biomass (circles) vs. A Predicted Growth (line)

Scenario: (*Model Refinement*) Certain resources (food, for instance) allow for the support of only a maximum population level (rather than one that increases without bound). As this maximum level is approached, the growth slows down. The data in Figure 2.4.5 show what actually happens to the yeast culture growing in a restricted area as time increases beyond the eight observations given in Figure 2.4.1.

Time in hours n	Observed yeast biomass p_n	Change in biomass $p_{n+1} - p_n$
0	9.6	
1	18.3	8.7
2	29.0	10.7
3	47.2	18.2
4	71.1	23.9
5	119.1	48.0
6	174.6	55.5
7	257.3	82.7
8	350.7	93.4
9	441.0	90.3
10	513.3	72.3
11	559.7	46.4
12	594.8	35.1
13	629.4	34.6
14	640.8	11.4
15	651.1	10.3
16	655.9	4.8
17	659.6	3.7
18	661.8	2.2

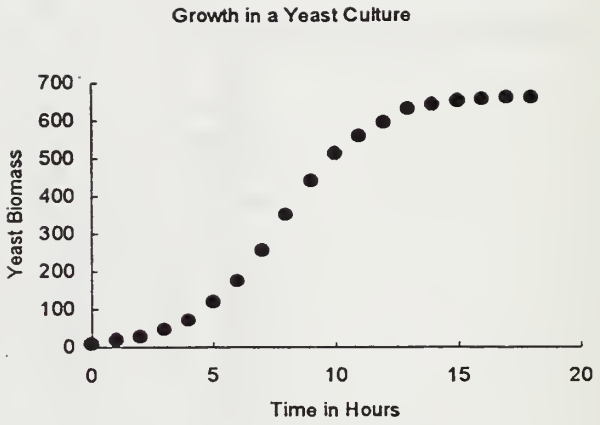


Figure 2.4.5: Yeast Biomass Approaches a Limiting Population Level

Using Excel: Here is where Microsoft Excel makes modeling change so easy. To create the data in Figure 2.4.5, just drag through the original worksheet data created previously for Figure 2.4.2, beginning with the last entry and continuing until Row 23 (where the time reaches 18 hours). The yeast biomass in the second column, still has to be entered by hand. Also, a new chart has to be created since we are graphing yeast biomass versus time instead of the change in yeast biomass versus the biomass itself, as in the previous chart exhibited in Figure 2.4.4. To clear the old chart from the worksheet (since it is no longer needed), simply select the chart and press **CLEAR** under the **EDIT** menu.

Scenario: (*continued*) Notice from the third column of the worksheet in Figure 2.4.5 that the change in population per hour becomes smaller as the resources become more limited or constrained. From the graph of population versus time (see Figure 2.4.5), the population appears to be approaching a limiting value or carrying capacity which we estimate to be about 665. As p_n approaches 665, the change slows considerably. Because $665 - p_n$ gets smaller as p_n approaches 665, we consider the logistic model (see Giordano, Weir, and Fox, op. cit., page 62):

$$\Delta p_n = p_{n+1} - p_n = k\,(665 - p_n)\,p_n\,.$$

Note that Δp_n becomes increasingly small as p_n approaches 665. Let us check this hypothetical model against the data and estimate the new constant k.

Using Excel: To create the data to test this model, highlight the second and third columns (with headings Observed Yeast Biomass and Change in Biomass per hour) from Figure 2.4.5. Select **COPY** from the **EDIT** menu, move to **Sheet2**, and **PASTE** the data in Columns A and B of Sheet2, beginning with cell A1. A new sheet is used here in order to keep the information and charts separated and organized. To create the data in the third column (Column C), type the following Excel formula into cell C6 adjacent to the first entry in Column B:

=A5*(665 - A5)

Highlight and drag this formula to fill Column C (see Figure 2.4.6).

	A	B	C	D
1	Observed	Change	Constrained	
2	yeast	in	Growth	
3	biomass	biomass	Model	
4	p_n	$p_{n+1} - p_n$	$p_n(665-p_n)$	
5	9.6			
6	18.3	8.7	6291.84	
7	29.0	10.7	11834.61	
8	47.2	18.2	18444.00	
9	71.1	23.9	29160.16	
10	119.1	48.0	42226.29	
11	174.6	55.5	65016.69	
12	257.3	82.7	85623.84	
13	350.7	93.4	104901.21	
14	441.0	90.3	110225.01	
15	513.3	72.3	98784.00	
16	559.7	46.4	77867.61	
17	594.8	35.1	58936.41	
18	629.4	34.6	41754.96	
19	640.8	11.4	22406.64	
20	651.1	10.3	15507.36	
21	655.9	4.8	9050.29	
22	659.6	3.7	5968.69	
23	661.8	2.2	3561.84	
24				

Figure 2.4.6: Data to Test the Hypothetical Model

Now create the chart, using the **(XY) Scatter** and Format **1** (one) as the chart type. One change is needed after the chart is generated in order to plot Column B versus Column C. Activate the chart, and highlight or select the data points. Choose the **SELECTED DATA SERIES** command under the **FORMAT** menu. A Format Data Series dialog box will appear. Select **X Values tab** and change the X values from column B to C. Then select the **Names and Values tab** and change the Y values from column C to B. Press the **OK** button. Therefore, the axes have been reversed (Excel chooses the data and axis by default, not in the order desired for the graph). Figure 2.4.7 shows the resulting chart.

Change in biomass $p_{n+1} - p_n$	Constrained Growth Model $p_n(665-p_n)$
8.7	6291.84
10.7	11834.61
18.2	18444.00
23.9	29160.16
48.0	42226.29
55.5	65016.69
82.7	85623.84
93.4	104901.21
90.3	110225.01
72.3	98784.00
46.4	77867.61
35.1	58936.41
34.6	41754.96
11.4	22406.64
10.3	15507.36
4.8	9050.29
3.7	5968.69
2.2	3561.84

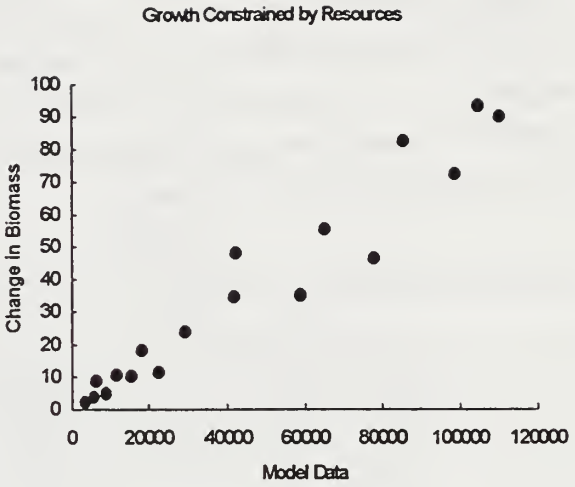


Figure 2.4.7: Testing the Constrained Growth Model

Scenario: (*Solve the Model Numerically and Verify the Results*) Examining Figure 2.4.7, the plot reasonably approximates a straight line projected through the origin. By accepting this proportionality argument, the slope of the line can be estimated as $k \approx 0.00082$ to give the model:

$$p_{n+1} - p_n = 0.00082\,(665 - p_n)\,p_n$$

Solving for p_{n+1} gives:

$$p_{n+1} = p_n + 0.00082\,(665 - p_n)\,p_n$$

Substitute $p_0 = 9.6$ into the expression to compute p_1:

$$\begin{aligned} p_1 &= p_0 + 0.00082\,(665 - p_0)\,p_0 \\ &= 9.6 + 0.00082\,(665 - 9.6)\,9.6 \\ &= 14.76 \end{aligned}$$

By iterating, to find the values of p_2, p_3, ... etc., a table of values can be computed to provide a numerical solution to the model. We plot the yeast biomass observations together with this numerical solution of the model predictions versus time on the same graph to see how good the model capture the *trend* of the observed data.

Using Excel: To create the data to plot this graph, first highlight Columns A and B from Sheet1 (which can be viewed in Figure 2.4.5) and **COPY** and **PASTE** the data in **Sheet3**. To create the data in the third

column (Column C), insert the initial value **9.6** in cell C5 (the cell adjacent to the first entry in Column B), and type the following Excel formula in cell C6 (the next cell below the input value 9.6):

=C5+0.00082*(665-C5)*C5

Now create the chart as before, but choose the data plots as open circles for the predicted values (Figure 2.4.8). As can be seen, the predictions capture the trend of the actual observed data.

Time in hours n	Observed Yeast Biomass p_n	Computed Model Predictions p_n
0	9.6	9.60
1	18.3	14.76
2	29.0	22.63
3	47.2	34.55
4	71.1	52.41
5	119.1	78.74
6	174.6	116.59
7	257.3	169.02
8	350.7	237.76
9	441.0	321.05
10	513.3	411.60
11	559.7	497.13
12	594.8	565.56
13	629.4	611.68
14	640.8	638.42
15	651.1	652.34
16	655.9	659.11
17	659.6	662.29
18	661.8	663.76

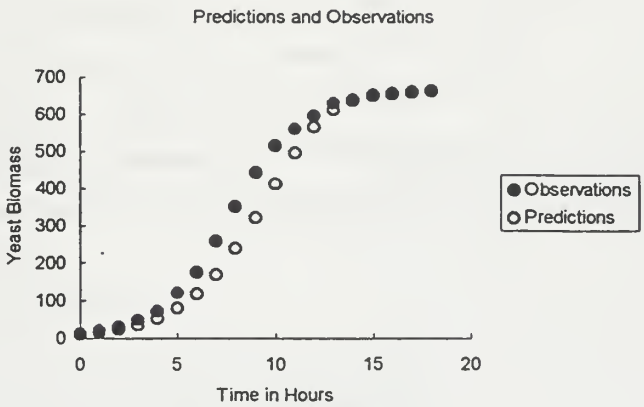


Figure 2.4.8: Model Predictions Together With the Observations.

E. NUMERICAL SOLUTIONS

In this section, numerical solutions to dynamical systems are built starting with an initial value and iterating a sufficient number of subsequent values to determine the patterns involved. The objective is to predict the future behavior and gain insight into how to influence the behavior, if necessary.

Example 2.5: Prescription for Digoxin

Scenario: Digoxin is used in the treatment of heart patients (see Giordano, Weir, and Fox, op. cit., Chapter 3.2, Problem 1, page 67). The objective here is to consider the decay of digoxin in the bloodstream in order to prescribe a dosage that keeps the concentration between acceptable (both safe and effective) levels. Suppose a daily drug dosage of 0.1 mg is prescribed, which results in the dynamical system:

$$a_{n+1} = 0.5a_n + 0.1$$

Let us consider three starting values, or initial doses

$$A: a_0 = 0.1,$$

$$B: a_0 = 0.2,$$

$$C: a_0 = 0.3$$

and compute the numerical solution for each initial dose.

Using Excel: This is an excellent example to demonstrate the power of Microsoft Excel. First, let Column A keep track of the number of iterations: in Column A, enter the values **0,1,...,15** beginning with cell A3 (only 15 iterations will be done at first). Then enter the starting values or initial doses, **0.1**, **0.2**, **0.3**, in cells B3, C3, and D3, respectively, and include the column titles or headings. Now enter the following Excel formula into cell B4, which will be used as a basis to calculate each iteration (Figure 2.5.1):

=0.5*B3 + 0.1

	A	B	C	D	E
1	Iterations	A	B	C	
2	*n*	a_n	a_n	a_n	
3	0	0.1	0.2	0.3	
4	1	=0.5*B3+0.1			
5	2				
6	3				
7	4				
8	5				
9	6				
10	7				

Figure 2.5.1: Example Diagram for the Digoxin Worksheet

Now there are two choices: either drag and copy the formula to cells C3 and D3 and then drag and copy all three cells (B3, C3, and D3) at the same time down to the end of Row 18, or drag and copy the formula from cell B3 to cell B18 and then drag and copy cells B3 through B18 (e.g., B3, B4, B5,...,B18) at the same time until the end of Column D. Either way, the formula is copied to all the cells in the table, and the data is correctly calculated for each cell in each column (i.e., for each starting value).

Finally, graph the data in each Column B-D on a single chart using three different symbols, using the number of iterations *n* for the x-axis values (see Figure 2.5.2). After the chart is generated, names to represent each data type in a legend can be typed by first highlighting each series of data points on the chart and choosing the **SELECTED DATA SERIES** command from the **FORMAT** menu. A Format Data Series dialog Box will appear. Select the **Name and Values tab**, and then type in the name to represent the data in the Name box. Repeat these steps for each sequence of data points.

Iterations	A	B	C
n	a_n	a_n	a_n
0	0.1	0.2	0.3
1	0.15	0.2	0.25
2	0.175	0.2	0.225
3	0.1875	0.2	0.2125
4	0.19375	0.2	0.20625
5	0.196875	0.2	0.203125
6	0.1984375	0.2	0.2015625
7	0.19921875	0.2	0.20078125
8	0.19960938	0.2	0.20039063
9	0.19980469	0.2	0.20019531
10	0.19990234	0.2	0.20009766
11	0.19995117	0.2	0.20004883
12	0.19997559	0.2	0.20002441
13	0.19998779	0.2	0.20001221
14	0.19999389	0.2	0.2000061
15	0.19999695	0.2	0.2000305

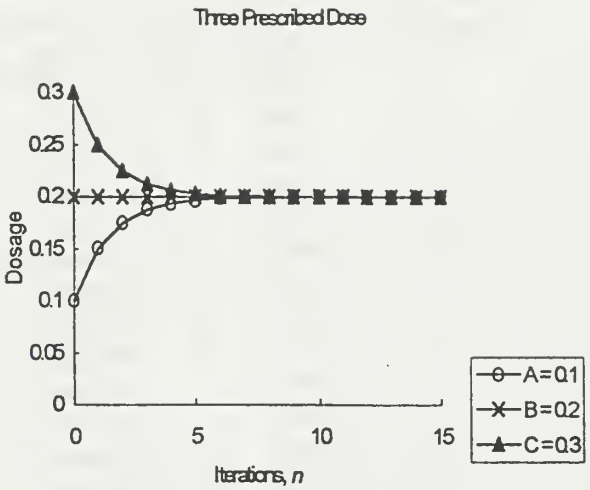


Figure 2.5.2: Three Initial Digoxin Doses

Scenario: (*continued*) Note that the value 0.2 is an equilibrium value because, once reached, the system remains at 0.2 forever. Further, the graph and numerical solutions show that if the initial values start above (as in Case C) or below (as in Case A) that value, then the equilibrium value is approached as a *limit*. Now, compute solutions for starting values even closer to 0.2, lending more evidence that 0.2 is a *stable* equilibrium value. When prescribing digoxin, the concentration level must stay above an effective level for a period of time without exceeding a safe level.

Using Excel: With Microsoft Excel, this will be easy! Go to the data table from Figure 2.5.2. **COPY** and **PASTE** the entire data table in **Sheet2**. Now, just change the starting values or initial doses located in cells B3 and D3 (leave 0.2 in cell C3 as a reference). Choose values closer to 0.2, such as **0.18** and **0.21**. Once these values are entered, Excel automatically updates the rest of the data to adjust the change (see Figure 2.5.3).

Iterations	A	B	C
n	a_n	a_n	a_n
0	0.18	0.2	0.21
1	0.19	0.2	0.205
2	0.195	0.2	0.2025
3	0.1975	0.2	0.20125
4	0.19875	0.2	0.200625
5	0.199375	0.2	0.2003125
6	0.1996875	0.2	0.20015625
7	0.19984375	0.2	0.20007813
8	0.19992188	0.2	0.20003906
9	0.19996094	0.2	0.20001953
10	0.19998047	0.2	0.20000977
11	0.19999023	0.2	0.20000488
12	0.19999512	0.2	0.20000244
13	0.19999756	0.2	0.20000122
14	0.19999878	0.2	0.20000061
15	0.19999939	0.2	0.20000031

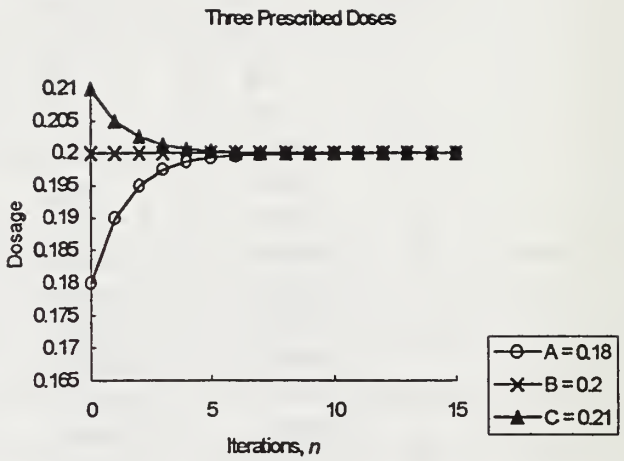


Figure 2.5.3: Solutions for Starting Values Closer to 0.2 Dosage

F. SYSTEMS OF DIFFERENCE EQUATIONS

For selected starting values, numerical solutions are built to get an indication of the long-term behavior of a system. For the systems considered in this section, the equilibrium values will be found. What happens near these values gives great insight concerning the long-term behavior of the system. The goal now is to model the behavior with difference equations and explore numerically the behavior predicted by the model.

Example 2.6: Competitive Hunter Models – Spotted Owls and Hawks

Scenario: In the spotted owl population, assume the existence of another species, the hawk, living in the same ecosystem. More specifically, assume the two species compete against each other for the available limited resources (such as food) in the habitat. The effect of the presence of the Hawk species is to diminish the growth rate of the spotted owl.

Let O_n represent the value of the spotted owl population at the end of day n, and let H_n denote the competing hawk population. Based on the competing species model, the following difference equations are obtained (see Giordano, Weir, and Fox, op. cit., page 86):

$$O_{n+1} = (1 + k_1) O_n - k_3 O_n H_n$$

$$H_{n+1} = (1 + k_2) H_n - k_4 O_n H_n$$

where k_1, k_2, k_3, and k_4 are positive constants.

Note that in the absence of the other species, each species exhibits a positive growth rate, and the presence of the second species diminishes the growth rate of the first. Now let us choose specific values for the constants of proportionality and consider the system:

$$O_{n+1} = 1.2\,O_n - 0.001\,O_nH_n$$

$$H_{n+1} = 1.3\,H_n - 0.002\,O_nH_n$$

If the equilibrium values are (O, H), then $O = O_{n+1} = O_n$ and $H = H_{n+1} = H_n$ simultaneously. Substituting into the system yields the simultaneous equations

$$O = 1.2\,O - 0.001\,OH$$

$$H = 1.3\,H - 0.002\,OH$$

or

$$0.2\,O - 0.001\,OH = O\,(0.2 - 0.001\,H) = 0 \qquad \text{(Eqn 1)}$$

$$0.3\,H - 0.002\,OH = H\,(0.3 - 0.002\,O) = 0 \qquad \text{(Eqn 2)}$$

The first equation indicates that there is no change in the owl population if $O = 0$ or $H = 0.2 \div 0.001 = 200$. The second equation indicates there is no change in the hawk population if $H = 0$ or $O = 0.3 \div 0.002 = 150$. Thus, equilibrium values exist at (O, H) = (0, 0) and (O, H) = (150, 200) because *neither* population changes at those points.

To analyze what happens in the vicinity of the equilibrium values, we build numerical solutions for the following three starting populations:

	Owls	Hawks
Case 1	151	199
Case 2	149	201
Case 3	10	10

Note that the first two values are close to the equilibrium value (150, 200), whereas the third is near the origin. Let us iterate these starting values using the above equations to obtain numerical solutions.

Using Excel: Let Column A record the number of iterations: Therefore, enter the values **0,1,...30** beginning with cell A2 (only 30 iterations will be done at first). Then enter the starting values, **151** owls and **199** hawks, in cells B2 and C2, respectively. Also, add column titles or headings, as necessary. Next enter the following Excel formula in cell B3 (which will be used as a basis to calculate each iteration in Column B, or the number of owls):

=1.2*B2 - 0.001*B2*C2

Likewise, enter the following Excel formula in cell C3 (which will be used as a basis to calculate each iteration in Column C, or the number of hawks):

=1.3*C2 - 0.002*B2*C2

Now drag each formula down their respective columns. Each formula is copied to all the cells in the table, and the data is correctly calculated for each cell in each column (e.g., for each starting value). Finally, graph the data on a chart using two different symbols versus the number of iterations, *n* in days (see Figure 2.6.1).

n	Owls	Hawks
1	151.00	199.00
2	151.15	198.60
3	151.36	198.14
4	151.64	197.60
5	152.01	196.96
6	152.47	196.17
7	153.05	195.20
8	153.79	194.00
9	154.71	192.53
10	155.87	190.72
11	157.31	188.48
12	159.12	185.73
13	161.40	182.34
14	164.25	178.18
15	167.83	173.10
16	172.34	166.93
17	178.04	159.47
18	185.26	150.53
19	194.42	139.91
20	206.11	127.48
21	221.05	113.18
22	240.25	97.09
23	264.97	79.57
24	296.88	61.27
25	338.06	43.27
26	391.05	27.00
27	458.70	13.98
28	544.03	5.35
29	649.92	1.13
30	779.17	0.00

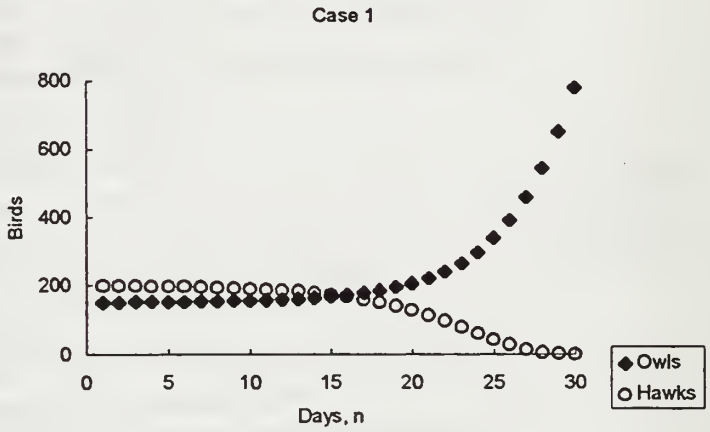


Figure 2.6.1: Owls versus Hawks for Case 1 Starting Values

The other two cases can be constructed just as easily. Go to the data table in Figure 2.6.1. For Case 2, **COPY** and **PASTE** the entire data table in **Sheet2**. For Case 3, **COPY** and **PASTE** the entire data table in **Sheet3**. Now, just change the starting values located in cells B2 and C2 in Sheet2 and Sheet3. Once these values are entered, Excel automatically updates the rest of the data to adjust the changes. These two cases, with corresponding figures can be found in Appendix C: Numerical Solutions to the Competitive Hunter Models.

Scenario: (*continued*) Suppose 350 owls and hawks are to be placed in a habitat modeled by the above equations. If 150 of the birds are owls, the model predicts the owls will remain at 150 forever. If one owl is removed from the habitat (e.g., 149), the model predicts that the owl population will die out. If 151 owls are placed in the habitat, however, the model predicts that the owls will grow without bound while the hawks disappear. Thus, the model is extremely sensitive to the initial conditions. Note that the model predicts that coexistence of the two species is impossible and that one species eventually dominates the habitat. Explore the system further by examining other starting points and by changing the coefficients of the model.

Using Excel: With Microsoft Excel, this will be easy. All that is required is to change the starting values located in cells B2 and C2 of Figure 2.6.1, or to change the constant coefficients in each Excel formula. Excel does all the calculating. It is a good idea to add columns instead of moving to a new sheet, in order to see how the changes in the starting values and the coefficients directly affect the data.

Having explored numerically several dynamical systems and examined the long-term behaviors of these systems, it is evident that the study of discrete change through the power of Excel can be modeled exactly by difference equations. Proportionality is also used to approximate change, and thus numerical solutions to the difference equations can be constructed by Excel to determine the types of long-term behaviors that they predict. Therefore, through Excel models are built to explain behavior and make predictions. Next, Excel formalizes the concept of proportionality and uses it to uncover relationships among variables selected in the model-building process.

III. MODELING USING PROPORTIONALITY

An important technique in constructing a model is by testing and applying proportionality arguments to a set of data. These constructs can form *submodels* of the model. Large data sets can be used to test graphically a proposed proportionality. If the test indicates the assumed proportionality is reasonable, then an initial estimate of the constant of proportionality can be made directly from the graph. A procedure for testing a proportionality submodel consists of the following steps:

1. Enter the data observed for the dependent and independent variables.
2. Create a scatterplot of the raw data points to check for trends, smoothness, and to identify potential data outliers.
3. Perform transformations of the data (if any) suggested by the submodel.
4. Plot the (possibly transformed) data to test for proportionality.
5. Estimate the constant of proportionality from the graph.

The concept of proportionality is useful to uncover relationships among some of the variables selected in the model-building process. The process of making graphical plots worthy of checking the proposed proportionality, and for estimating the constants of proportionality using Excel, is presented here. The goal is to use Excel to gain a visual, qualitative estimate of the worthiness of a submodel, and to help in estimating the constants of proportionality.

So far, only the most essential skills have been discussed that enable the user to work through a basic mathematical modeling process using Microsoft Excel. This chapter expands on that working knowledge to help with more complicated, less familiar techniques. In order to simplify formulas (which saves time and enhances Excel's capabilities), cell reference operators and built-in functions are first presented, along with an introduction to the use of *Function Wizard*. Additionally, two new Microsoft Excel tools are introduced to help analyze the data on a chart. These include adding a *trendline* or *error bars* to a data plot. Trendlines are commonly used to study problems of prediction with a model. Error bars express the error factor visually and show the amount of error relative to some data marker.

A. ADVANCED SKILLS FOR BUILDING WORKSHEETS

This section provides advanced skills for working with cell references to allow the user to carry out operations involving sophisticated functions. Although many modeling examples presented in this thesis can be done without the material to follow, the suggestions presented here enhance and simplify techniques required for the entire mathematical modeling process.

1. Cell Reference Operators

There are three types of reference operators: *range*, *union*, and *intersection*. The range is designated by a ***colon*** (**:**) and produces one reference to all the cells between and including the two references. For example, B2:D2 denotes the range represented by the cells B2, C2, and D2; A1:A5 denotes the range represented by the cells A1, A2, A3, A4, and A5. The union is designated by a ***coma*** (**,**) and produces one reference to include both references. For example B4,D4 represents the union of cells B4 and D4. The intersection is designated by a ***space*** ( ) and produces one reference to cells common to the two references. For example, A1:A4 A3:A5 is A3 and A4, the intersection of {A1, A2, A3, A4} with {A3, A4, A5}. When using a range reference operator to refer to entire columns, rows (or a range of entire columns or rows), use the following abbreviated forms of the references:

Reference	Input into Formula
All of column A	A:A
All of columns A through F	A:F
All of row 1	1:1
All of rows 1 through 5	1:5
Entire worksheet	A:IV or 1:16384

2. Built-in Functions

Cell references can be grouped and Excel provides built-in formulas or *functions* to speed up the function formulation process. Functions can be used alone or as building blocks in larger formulas. The values given to a function on which to perform its operations are called *arguments*. Some of the familiar, common functions with their appropriate meaning are:

Function	Definition
ABS(*number*)	Returns the absolute value of a number
AVERAGE(*number1, number2, ...*)	Returns the average of its arguments
COS(*number*)	Returns the cosine of a real number
COUNT(*value1, value2, ...*)	Counts how many numbers are in the list of arguments
DEGREES(*angle*)	Converts radians to degrees
EXP(*number*)	Returns *e* raised to the power of a given number
FACT(*number*)	Returns the factorial of a number
LN(*positive number*)	Returns the natural logarithm of a number
MAX(*number1, number2, ...*)	Returns the maximum value in a list of arguments
MIN(*number1, number2, ...*)	Returns the minimum value in a list of arguments
PI( )	Returns the value of Pi
POWER(*number, power*)	Returns the result of a number raised to a power
PRODUCT(*number1, number2, ...*)	Multiplies its arguments
QUOTIENT(*numerator, denominator*)	Returns the integer portion of a division
RADIANS(*angle*)	Converts degrees to radians
RAND( )	Returns a random number between 0 and 1
ROUND(*number, num_digits*)	Rounds a number to a specified number of digits
SIN(*number*)	Returns the sine of a real number

Function	Definition
SQRT(*positive number*)	Returns a positive square root
SUM(*number1, number2, ...*)	Adds its arguments
TAN(*number*)	Returns the tangent of a real number

To view the entire list of Excel built-in functions, choose the **FUNCTION** command under the **INSERT** menu, which brings up the Function Wizard dialog box. The user can choose a Function Category as well as a Function Name. To exit the Function Wizard dialog box, press Cancel.

Functions are used by entering them in formulas. For example, instead of typing the formula '=A1+A2+A3+A4' , the SUM function can be used to build the formula '=SUM(A1:A4)', which also uses the range of a cell reference operator to simplify the formula. The sequence of characters used to enter a valid function is called the *syntax*. All functions have the same basic syntax. Parentheses are used to enclose all the arguments, and commas are used to separate individual arguments within the parentheses. If this syntax is not followed, Excel displays a message indicating that there is an error in the formula. Also, Excel requires a function name to be in uppercase letters when the formula is entered (a function typed with lowercase letters is not valid). The illustration that follows is an example of a function and its syntax (Figure 3-A.2.1).

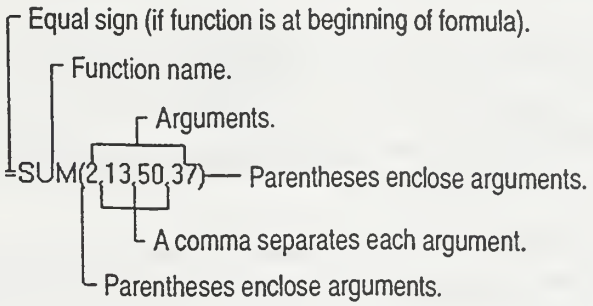


Figure 3-A.2.1: Example of a Function and its Syntax

In order to use the functions, the following guidelines must be followed. Parentheses tell Excel where the arguments begin and end. Remember to include both left and right parentheses with no preceding or following spaces. Do not use commas to separate thousands for numeric values, (e.g., 1,254,907 is incorrect, but 1254907 is valid as an argument). Arguments can consist of numbers, references, text, logical values, arrays, error values, constant values, or formulas. If a formula is used as an argument, it can contain other functions as well. When an argument to a function is itself a function, it is said to be *nested*. In Microsoft Excel, functions can be nested up to seven (7) levels in a formula. An example is given in the next subsection.

Finally, the SUM function is the most frequently used of all worksheet functions. Even more convenient than the SUM function is the **AutoSum** icon (see Figure 3-A.2.2) on the Standard toolbar.

Figure 3-A.2.2: The AutoSum Icon Located on the Toolbar

When the AutoSum icon is used, Excel automatically types the SUM function and even suggests the range of cells to be added. To enter a sum formula, select a cell adjacent to a row or column of numbers to be added, and then click the **AutoSum** icon. This activated cell will contain the results of the SUM function. A flashing *marquee* (dashed line) will enclose the cells to be summed. If the suggested range is incorrect, or if Excel cannot determine a range to suggest, simply drag through the correct range and press **Enter** to accept the completed formula.

3. Use of the Function Wizard

Although a function can be typed directly into a formula, the easiest method of inserting functions is to use the Function Wizard. To use a built-in Microsoft Excel function, the Function Wizard can be used to help select a function, to assemble the arguments correctly, and to insert the function into the formula. The formula bar shows the changes made while building the formula. To add a function to a formula, activate the Function Wizard by clicking the **Function Wizard** icon (see Figure 3-A.3.1) located in the toolbar. A two-step Function Wizard dialog box will appear.

Figure 3-A.3.1: The Function Wizard Icon Located on the Toolbar

Clicking the **Finish** button in the Step 1 dialog box automatically inserts the selected function into the formula with the argument names inserted as place holders. The argument text can then be replaced with the values necessary to complete the function. To go from Step 1 to the Step 2 dialog box, click the **Next** button which displays the Step 2 dialog box. Here is where numbers, references, names, formulas, text, or other functions can be entered into the argument edit boxes. After valid values are entered for each required argument, the calculated value for the function appears in the Value box located atc the top, right corner of the Step 2 dialog box. Click the **Finish** button to insert the completed function into the formula.

To enter functions as arguments to other functions, as in the formula =ABS(AVERAGE(B4,SUM(D4:D8))), is considered a nested function. This formula has two levels of nested functions. The SUM function is entered as an argument to the AVERAGE function, which is itself an argument to the ABS function. To nest a function, click the small **Function Wizard** button in the appropriate argument edit box in the Function Wizard Step 2 dialog box. When this button is entered, another Function Wizard dialog box appears, allowing the user to nest another function as an argument.

Finally, to edit functions in an existing formula, select a cell containing an existing formula that includes functions, and choose the **Function Wizard** icon. The first function in the formula is opened in the Function Wizard Editing mode (see Figure 3-A.3.2), allowing the user to modify the arguments. This is also an excellent way to debug a function. For example, if the user clicks the **Function Wizard** icon while a cell containing the formula =SUM(ABS(B2),C2,D2) is selected, the following dialog box appears (Figure 3-A.3.2).

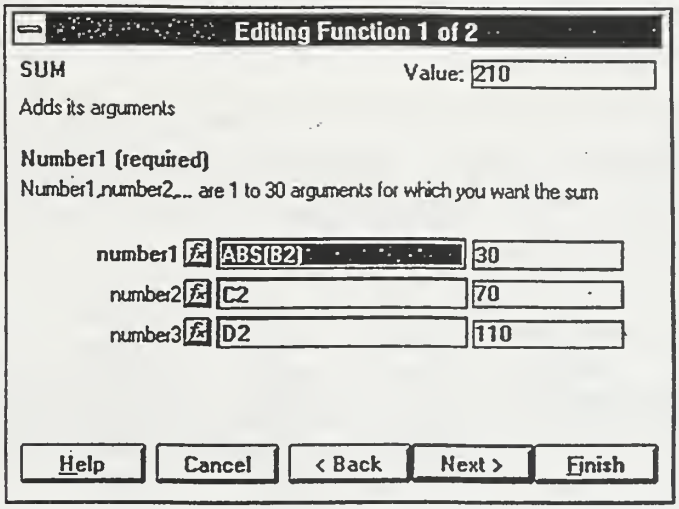


Figure 3-A.3.2: Example of the Function Wizard Editing Mode

The Function Wizard opens the first function in the editing version of the Step 2 dialog box in order to edit the arguments. Notice that the dialog box title is Editing Function 1 of 2. By clicking the Next button, the changes are entered for the current function, and the second function in the formula (e.g., the ABS function) is opened for editing.

B. VEHICULAR STOPPING DISTANCE

A popular "Rule of Thumb" often given to students in driver education classes is the "Two Second Rule" to prescribe a safe following distance. The rule states that if a driver stays two seconds behind the car in front, then the driver has the correct distance no matter what the speed. Since the amount of time is constant (2 seconds), the rule suggests a proportionality between stopping distance and speed. To test this rule, the following problem is posed.

Example 3.1: Vehicle Stopping Distance

Scenario: Let us predict a vehicle's total stopping distance as a function of time. In the development of this model (see Giordano, Weir and Fox, op. cit., page 103), total stopping distance is calculated as the sum of the reaction distance d_r and braking distance d_b. The following submodels are

hypothesized in that development:

$$d_r \propto v$$

$$d_b \propto v^2$$

to give the total stopping distance

$$d = k_1 v + k_2 v^2$$

where k_1 and k_2 are constants.

At this point, the submodels could be tested against the data provided by the U.S. Bureau of Public Roads given in Table 3.1.1. To test the submodel for reaction distance, plot the driver reaction distance against velocity (see Figure 3.1.1). The data should lie approximately on a straight line passing through the origin.

Speed (mph)	Driver Reaction Distance (ft)	Braking Distance (ft)	Average Braking Distance (ft)	Observed Total Stopping Distance (ft)	Average Stopping Distance (ft)
20	22	18	20	40-44	42
25	28	25	28	53-59	56
30	33	36	40.5	69-78	73.5
35	39	47	52.5	86-97	91.5
40	44	64	72	108-124	116
45	50	82	92.5	132-153	142.5
50	55	105	118	160-186	173
55	61	132	148.5	193-226	209.5
60	66	162	182	228-268	248
65	72	196	220.5	268-317	292.5
70	77	237	266	314-372	343
75	83	283	318	366-436	401
80	88	334	376	422-506	464

Table 3.1.1: Observed Reaction and Braking Distances

Similarly, to test the submodel for braking distance, plot the observed braking distance against the velocity squared. Proportionality seems to be a reasonable assumption at the lower speed, although it does seem to be less convincing at the higher speeds (see Figure 3.1.2).

Using Excel: Entering the data from the table onto an Excel worksheet is routine, as well as generating the charts for that data, as seen in Figures 3.1.1 and 3.1.2.

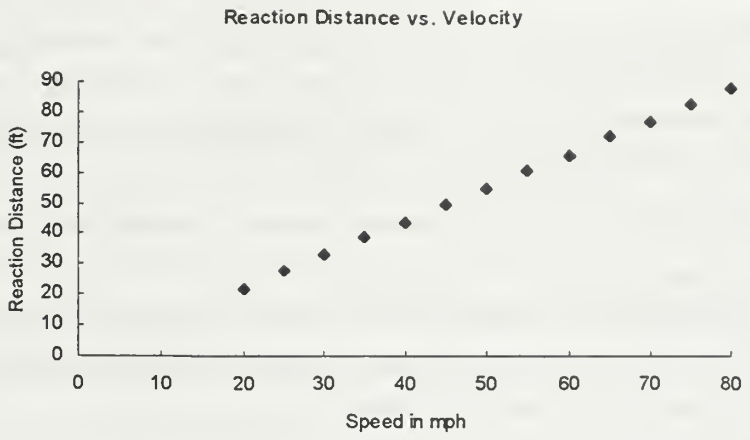


Figure 3.1.1: Driver Reaction Distance vs. Velocity

However, one additional column of data is needed in order to graph Figure 3.1.2. This column squares the speed or velocity using the following Excel formula:

=SUMSQ(A4)

The Function Name *SUMSQ* represents the sum of the squares of the argument and is obtained from the **Function Wizard** under the **Math & Trig** Function Category listing. To use the Function Wizard, first highlight the cell in which the formula is to be entered, then press the **Function Wizard** icon.

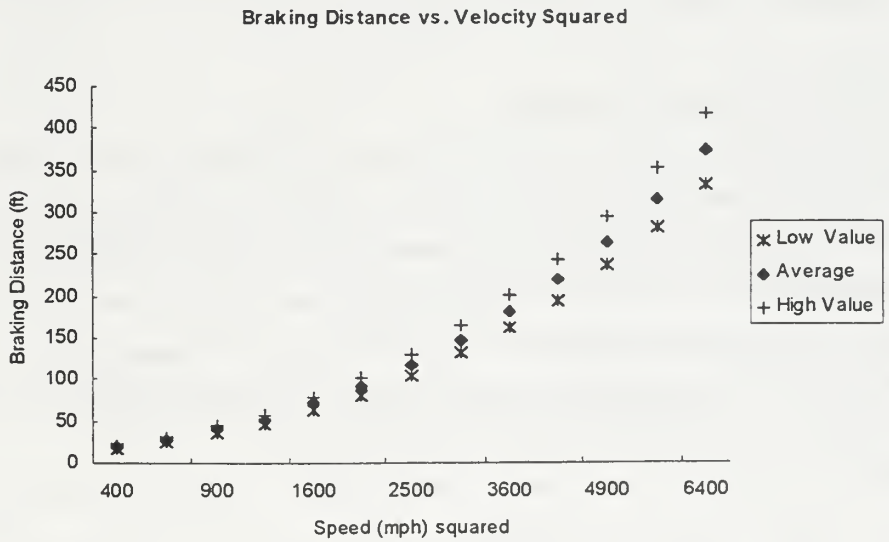


Figure 3.1.2: Observed Braking Distance vs. Velocity Squared

Once the Function Wizard dialog box appears, find the Function Name under the appropriate category, and then double-click on that name. Excel then asks for the number (e.g., cell location for that number, in this case A4), either enter it from the keyboard or go to it in the worksheet, then press **Finish** to execute the formula. Now drag this formula down the entire column. Use this new column of data (e.g., velocity squared) as the interval for the x-axis in Figure 3.1.2. Plot the high value, the low value, and the average value for braking distance as provided in Table 3.1.1.

Scenario: (*continued*) For the submodel which plots the driver reaction distance against velocity (see Figure 3.1.1), the graph appears to be a straight line of approximate slope 1.1 passing through the origin (see Figure 3.1.3). Therefore, the following submodel is obtained:

$$d_r = 1.1v$$

For the submodel which plots the observed braking distance against the velocity squared, proportionality seems to be a reasonable assumption at lower speeds, although it does seem to be less convincing at higher speeds (see Figure 3.1.2). By graphically fitting a straight line (by hand) to the data for average braking distance (see Figure 3.1.4), the slope is estimated to be 0.054 and the following submodel is obtained:

$$d_b = 0.054\, v^2$$

Connecting the two submodels, yields the following model for the total stopping distance

$$d = 1.1v + 0.054\, v^2$$

Using Excel: Now, however, to graph these straight lines on the charts, new information regarding Excel is needed to complete the graphs in order to predict their slopes.

1. Adding a Trendline to a Data Series

The first tool is adding a *trendline* to a data series in a chart to show the trend, or the direction, of the data in the series. Trendlines can be added to xy (scatter) charts. The first step in creating a trendline is to select the data series on the chart that the trendline is to be associated with. Then choose the **TRENDLINE** command from the **INSERT** menu. Select the type of trendline from the **Type tab**, and choose the options from the **Options tab**. After creating a trendline, its color, style, and weight can be changed by double-clicking the trendline on the chart to display the Format Trendline dialog box, or by simply highlighting the trendline and entering the **SELECTED TRENDLINE** from the **FORMAT** menu.

The following are the five types of trendlines to choose from:

Type	Description
Linear	Creates the trendline using the linear equation $y = mx + b$
Logarithmic	Creates the trendline using the logarithmic equation $y = c\, lnx + b$

Type	Description
Polynomial	Creates the trendline using the polynomial equation $y = b + c_1x + c_2x^2 + ... + c_6x^6$
Power	Creates the trendline using the power equation $y = cx^b$
Exponential	Creates the trendline using the exponential equation $y = ce^{bx}$

The trendline can be used to forecast forward, backward, or both, for the number of periods specified under the Options tab. Also, additional information can be displayed on the chart, including the trend line equation with or without the y-intercept. Excel automatically calculates these equations based on the trendline itself. These equations can be moved and formatted on the chart just like other data labels.

Using Excel: (*continued*) Create the charts for the data as usual, but to create the trendline for Figure 3.1.3, highlight or select the data points on the chart. Then choose **SELECTED TRENDLINE** from the **INSERT** menu, in which the Format Trendline dialog box appears. Select the **Type tab** and choose **Linear** from the Trend/Regression Type list, then from the **Options tab**, activate **Set Intercept = 0** (in order for the trendline to go through the origin) and activate **Display Equation on Chart**. Also, change the **Forward** and **Backward** Forecast from 0 to **0.5** Periods in order for the trendline to cover the length of the x-axis.

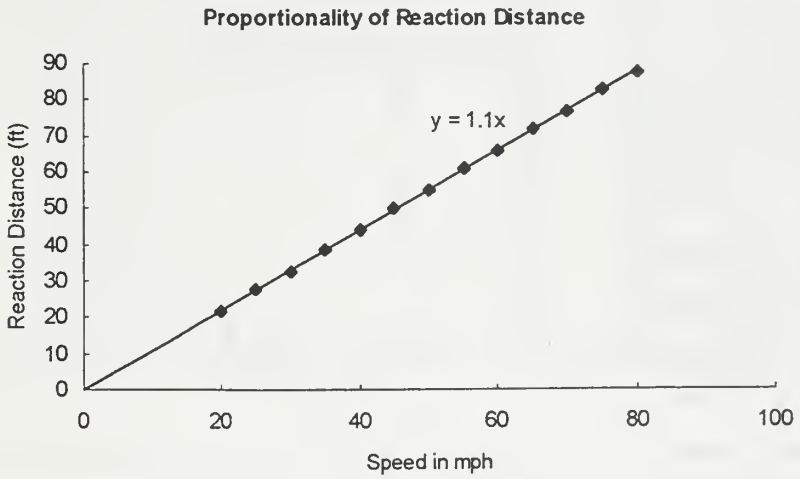


Figure 3.1.3: Proportionality of Reaction Distance and Speed

2. Adding Error Bars to a Data Series

The second new Microsoft Excel tool to introduce is adding *error bars* to a data series in a chart to indicate the degree of uncertainty -- the "plus or minus range" -- for each data point in the series. After the error bars are added to the data series, they remain associated with that series. If the series is moved, for example, by changing

the plot order, the error bars move with it. If the data in a series is changed, the error bars are recalculated and adjusted accordingly.

The first step in creating error bars is to select the data series on the chart for the error bars to be associated with. Then choose the **ERROR BARS** command from the **INSERT** menu, select the type of display, and specify how the error amount should be calculated or obtained. The error amount can be specified from the following options: Fixed, Percentage, Standard Deviation, Standard Error, and Custom. After creating the error bars, their color, style, and weight can be changed by double-clicking one error bar to display the Format Error Bars dialog box. The formatting changes will be applied to all the error bars for that data series.

Using Excel: (*continued*) For Figure 3.1.4, create the trendline as before (using the data points for the average braking distance in Table 3.1.1). There are two ways to create the error bars for this particular problem. First, simply graph the data points for the average braking distance, and then select each data point one at a time, choose **SELECTED ERROR BARS** from the **INSERT** menu, and then follow the choices in the Format Error Bars dialog box. However, be sure to use the **Custom** Error Amount and input the high and low braking distance from the table for each point. The other method is to plot all three values for the braking distance (e.g., high, low, and average) on a **Line** chart, choosing Format **7**, which is the high-low lines chart (e.g., it automatically adds the lines between the data points). The chart is then complete.

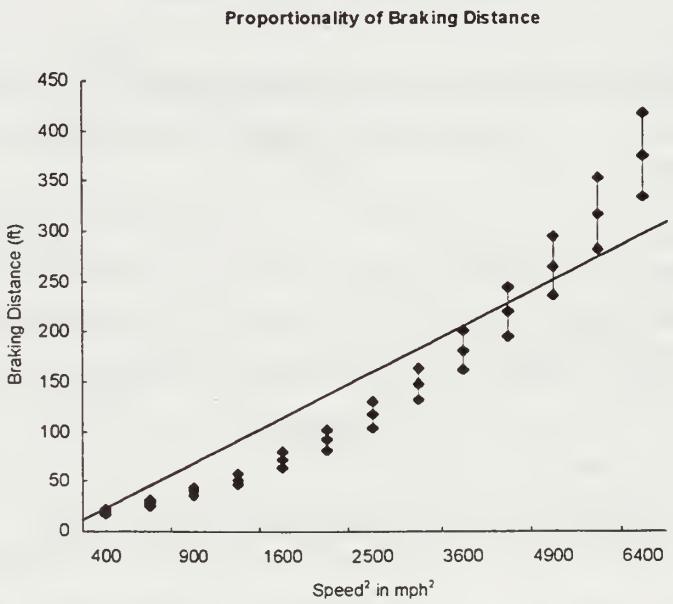


Figure 3.1.4: Proportionality of Braking Distance and the Speed Squared

Scenario: (*continued*) The predictions of the model for the total stopping distance recorded in Table 3.1.1 are plotted in Figure 3.1.6. Considering the grossness of the assumptions and inaccuracies of the data, the model seems to agree fairly reasonably with the observations up to about 80 mph. The rule of one 15-ft car length for every 10 mph of speed is also plotted in Figure 3.1.6, the data is computed in Column I of Figure 3.1.5. The rule significantly under estimates the total stopping distance at speeds exceeding 40 mph (see Giordano, Weir, and Fox, op. cit., page 107).

Using Excel: In order to plot Figure 3.1.6, two Excel formulas must be used to create a two new columns of data. To create the Predicted Braking Distance based on the model for the total stopping distance, use the following Excel formula (based on the data as it appears in Figure 3.1.5):

=1.1*A4+0.054*B4

and enter it in cell H4 a s depicted in Figure 3.1.5. To create the Braking Distance according to the rule "every 10 mph equals 15-ft of braking distance", use the following Excel formula:

=15*(A4/10)

and enter it in cell I4, as depicted in Figure 3.1.5.

	A	B	C	D	E	F	G	H	I
1			Driver Reaction	Braking	Average	Observed	Average	Predicted	"Rule"
2	Speed	Speed^2	Distance	Distance	Braking	Total Stopping	Stopping	Stopping	Stopping
3	(mph)	(ft^2)	(ft)	(ft)	Distance (ft)	Distance (ft)	Distance (ft)	Distance	Distance
4	20	400	22	18	20	40-44	42	43.6	30
5	25	625	28	25	28	53-59	56	61.25	37.5
6	30	900	33	36	40.5	69-78	73.5	81.6	45
7	35	1225	39	47	52.5	86-97	91.5	104.65	52.5
8	40	1600	44	64	72	108-124	116	130.4	60
9	45	2025	50	82	92.5	132-153	142.5	158.85	67.5
10	50	2500	55	105	118	160-186	173	190	75
11	55	3025	61	132	148.5	193-226	209.5	223.85	82.5
12	60	3600	66	162	182	228-268	248	260.4	90
13	65	4225	72	196	220.5	268-317	292.5	299.65	97.5
14	70	4900	77	237	266	314-372	343	341.6	105
15	75	5625	83	283	318	366-436	401	386.25	112.5
16	80	6400	88	334	376	422-506	464	433.6	120

Figure 3.1.5: An Example of an Excel Worksheet to Create the Stopping Distance Chart

When constructing the chart (see Figure 3.1.6), choose the **XY (Scatter)** as the type of chart with Format **1** (one). Then after the chart is graphed, select the data series representing the General Braking Distance and change the format from a Marker to a **Line** under the **Pattern tab** in the Format Data Series dialog box.

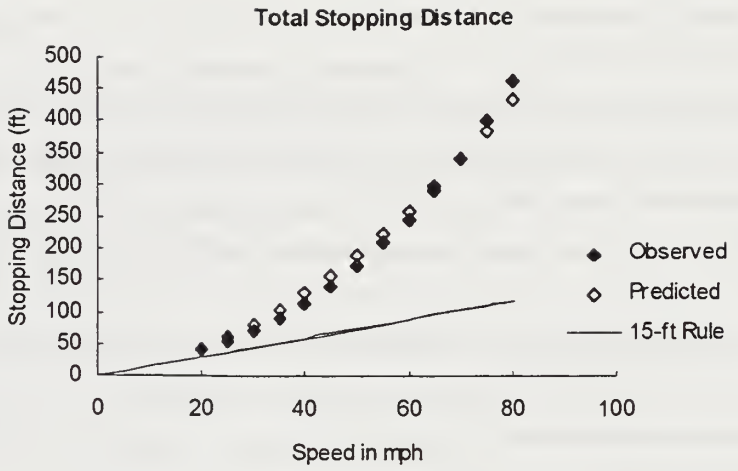


Figure 3.1.6: Total Stopping Distance

The predicted model for stopping distance and the actual observed stopping distance are plotted in Figure 3.1.6. The model seems to agree fairly reasonably with the observations up to about 80 mph. The rule of one 15-ft car length for every 10 mph of speed, which is also plotted in Figure 3.1.6, significantly underestimates the total stopping distance at speeds exceeding approximately 25 mph. Therefore, if the driver of the trailing vehicle must be fully stopped by the time he or she reaches the point occupied by the lead vehicle at the exact time of the observation, then the driver must trail the lead vehicle by the total stopping distance, either based on the predicted model or on the observed data themselves.

C. A BASS FISHING DERBY

Consider a sport fishing club that for conversation purposes wishes to encourage its membership to release their fish immediately after catching them. Therefore, how does someone fishing determine the weight of a fish he or she has caught? One might suggest that each individual carry a small portable scale. However, portable scales tend to be inconvenient and inaccurate, especially for smaller fish. The problem of predicting the weight of a fish in terms of some easily measurable dimensions is identified as follows. Because a general rule for sport fishing is sought, let us initially restrict attention to a single species of fish, say bass, and assume that within the species the average weight density is constant.

Example 3.2: A Bass Fishing Derby

Scenario: In the example of the bass fishing derby (see Giordano, Weir, and Fox, op. cit., page 121), let us proportion weight, W, to length, l, of a fish in order to test the model

$$W \propto l^3$$

where the length of the fish is chosen as the characteristic dimension.

First, consider the following data collected during a fishing derby:

Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
Weight, W (oz.)	27	17	41	26	17	49	23	16

Table 3.2.1: Length versus Weight for Several Bass

If the model is correct, the graph of W versus l^3 should approximate a straight line passing through the origin.

Using Excel: Entering the data from Table 3.2.1 onto an Excel worksheet is routine, however, an additional row of data is required in order to graph Figure 3.2.2. This additional row cubes the length (given in the first row) of each bass using the following Excel formula:

=(B1)^3

The results can be found in Figure 3.2.1.

	A	B	C	D	E	F	G	H	I	J
1	Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625	
2	Weight, W (oz.)	27	17	41	26	17	49	23	16	
3										
4	Volume, l^3 (in^3)	3048.63	1953.13	5132.95	3048.63	2012.31	5592.36	2818.16	2012.31	
5										

Figure 3.2.1: Length, Weight, and Volume for Several Bass

Now, create the chart for the data as usual, plotting the weight (W) versus the volume (l^3). To create the trendline for Figure 3.2.2, highlight or select the data points on the newly created chart. Then choose **SELECTED TRENDLINE** from the **INSERT** menu, in which the Format Trendline dialog box appears. Select the **Type tab** and choose **Linear** from the Trend/Regression Type list, then from the **Options tab**, activate **Set Intercept = 0** (in order for the trendline to go through the origin) and activate **Display Equation on Chart**. Also, change the **Backward** Forecast from 0 to **1950** Units in order for the trendline to cover the length of the x-axis (see Figure 3.2.2).

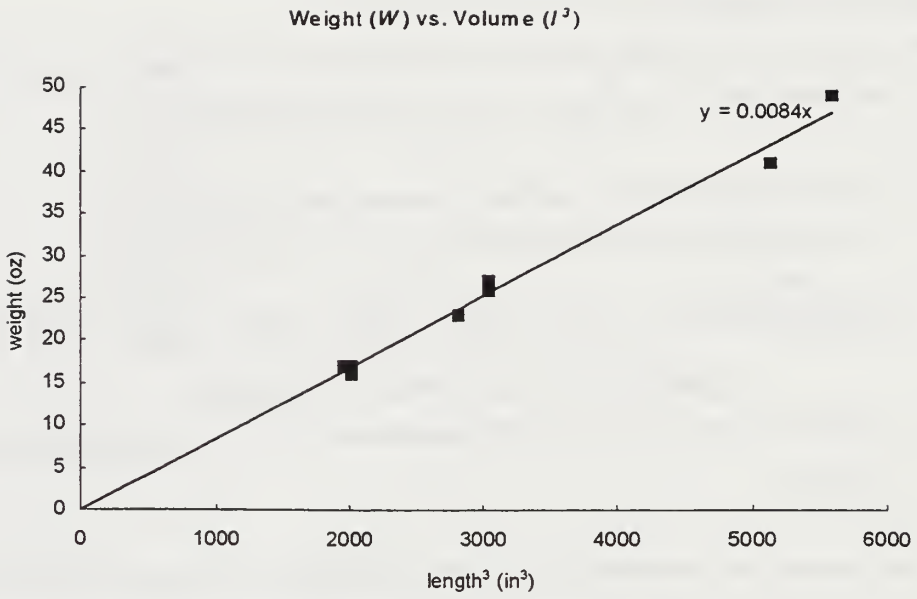


Figure 3.2.2: Graph of Weight (W) versus Volume (l^3) with a Trendline

Scenario: (*continued*) Now graph weight versus length using the model $W = 0.0084\ l^3$ (which Excel derived from the slope of the plot), showing also a plot of the original data points.

Using Excel: To graph the model $W = 0.0084\ l^3$, use the following Excel formula to find the data points for weight based on the cubed length of the bass:

=0.0084*((B1)^3)

Notice how close the model or estimated weight is to the actual weight of each bass in Figure 3.2.3.

	A	B	C	D	E	F	G	H	I
1	Length, *l* (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
2	Weight, *W* (oz.)	27	17	41	26	17	49	23	16
3									
4	Weight, W=0.0084*l^3	25.61	16.41	43.12	25.61	16.90	46.98	23.67	16.90

Figure 3.2.3: Data for the Model, $W = 0.0084\ l^3$

To plot the weight versus the length using the model, click on the **ChartWizard** icon and follow the 5 step process in the ChartWizard dialog box. In Step 1, since the data is located in non-adjacent rows, highlight Row 1, columns B through I, then insert a comma (,) after this range in the dialog box window. Now highlight the other set of data in Row 4, Columns B through I. Therefore, both sets of ranges appear in the first window of the ChartWizard dialog box. Even though the data is

plotted as a continuous curve, select the chart type **(XY) Scatter** with format **1** (one). The (XY) scatterplot is chosen in order to change the x-axis scale.

Now choose **SELECTED TRENDLINE** from the **INSERT** menu, in which the Format Trendline dialog box appears. Select the **Type tab** and choose **Power** from the Trend/Regression Type list, then from the **Options tab**, activate **Set Intercept = 0** (in order for the trendline to go through the origin) and activate **Display Equation on Chart**. Also, change the **Backward** Forecast from 0 to **12.5** Units in order for the trendline to cover the length of the x-axis.

Now that the trendline is inserted, highlight the data points directly on the chart. Choose the **SELECTED DATA SERIES** option under the **FORMAT** menu, and then when the Format Data Series dialog box appears, select the **Patterns tab**. Change the **Marker** selection to **None** which erases the data markers, but the trendline remains.

Once the curve for the model is plotted, add the original data points for the actual weight of each bass to the chart by activating the chart, and then selecting **NEW DATA** under the **INSERT** command. The new range of data points is automatically added to the chart. Some editing is required to obtain a plot exactly like the one presented here (see Figure 3.2.4).

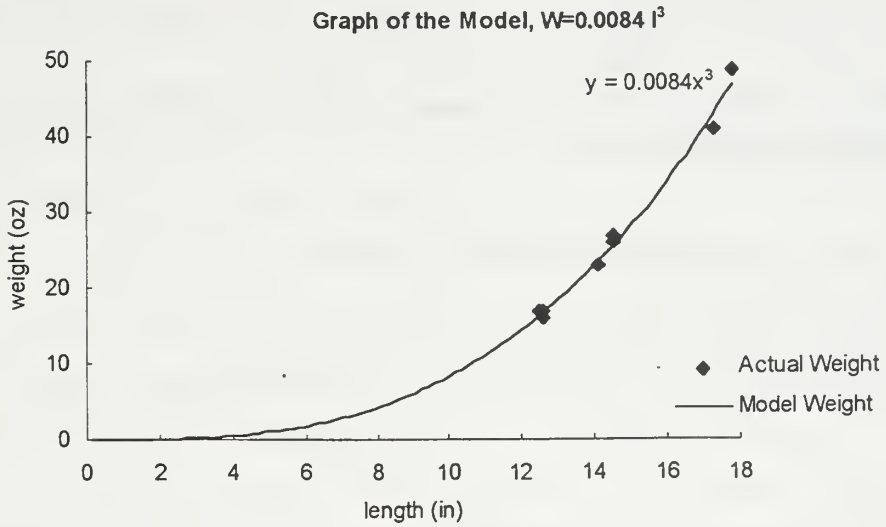


Figure 3.2.4: Graph of the Model, $W = 0.0084\ l^3$

Scenario: (*continued*) As suggested in Giordano, et. al., let us test the model

$$W = klg^2$$

for some positive constant k, (which is probably more satisfying to a fisherman), and where weight is designated by W, length is designated by l, and girth (which is the circumference of a fish at its widest point) is designated by g.

To get an initial test of the model, consider the following data:

Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
Girth, g (in.)	9.75	8.375	11	9.75	8.5	12.5	9	8.5
Weight, W (oz.)	27	17	41	26	17	49	23	16

Table 3.2.2: Length, Girth, and Weight of Several Bass

Because the model suggests a proportionality between W and lg^2, consider a plot of W versus lg^2.

Using Excel: Entering the data from Table 3.2.2 onto an Excel worksheet is routine, however, an additional row of data is required in order to graph Figure 3.2.6. This additional row is the estimated data for the weight given in the model, $W \propto lg^2$. The new variable lg^2 is found using the following Excel formula:

=B1*(B2)^2

The results can be found in Figure 3.2.5.

	A	B	C	D	E	F	G	H	I
1	Length, l (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
2	Girth, g (in.)	9.75	8.375	11	9.75	8.5	12.5	9	8.5
3	Weight, W (oz.)	27	17	41	26	17	49	23	16
4									
5	lg^2	1378.41	876.76	2087.25	1378.41	912.16	2773.44	1144.13	912.16

Figure 3.2.5: Length, Girth, Weight

Now graph the actual weight (W) versus the proportionality (lg^2) and include a trendline that passes through the origin (see Figure 3.2.6).

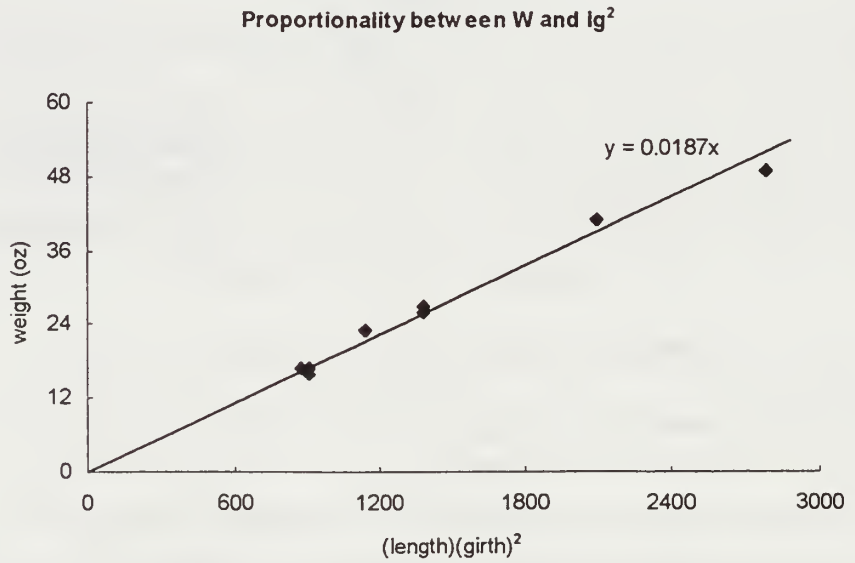


Figure 3.2.6: Graph of the Proportionality Between the Actual Weight (W) and lg^2

Using Excel: (*continued*) Excel has calculated the slope of the trendline. This slope is the estimate for k in $W = klg^2$. Therefore, this leads to the model:

$$W = 0.0187\ lg^2$$

To graph the model $W = 0.0187\ lg^2$, use the following Excel formula to find the data points for weight based on the length times the girth squared of the bass:

=0.0187*((B1)^3)

Notice how close the model or estimated weight is to the actual weight of each bass in Figure 3.2.7.

	A	B	C	D	E	F	G	H	I
1	Length, *l* (in.)	14.5	12.5	17.25	14.5	12.625	17.75	14.125	12.625
2	Girth, g (in.)	9.75	8.375	11	9.75	8.5	12.5	9	8.5
3	Weight, *W* (oz.)	27	17	41	26	17	49	23	16
4									
5	Weight, W=0.0187lg^2	25.78	16.40	39.03	25.78	17.06	51.86	21.40	17.06

Figure 3.2.7: Data for the Model, $W = 0.0187\ lg^2$

To plot the weight versus the length using the model, click on the **ChartWizard** icon and follow the 5 step process in the ChartWizard dialog box. In Step 1, since the data is located in non-adjacent rows, highlight Row 1, columns B through I, then insert a comma (,) after this range in the dialog box window. Now highlight the other set of data in Row 4, Columns B through I. Therefore, both sets of ranges appear in the first window of the ChartWizard dialog box. Even though the data is plotted as a continuous curve, select the chart type **(XY) Scatter** with format **1** (one). The (XY) scatterplot is chosen in order to change the x-axis scale.

Now choose **SELECTED TRENDLINE** from the **INSERT** menu, in which the Format Trendline dialog box appears. Select the **Type tab** and choose **Power** from the Trend/Regression Type list, then from the **Options tab**, activate **Set Intercept = 0** (in order for the trendline to go through the origin) and activate **Display Equation on Chart**. Also, change the **Backward** Forecast from 0 to **12.5** Units in order for the trendline to cover the length of the x-axis.

Now that the trendline is inserted, highlight the data points directly on the chart. Choose the **SELECTED DATA SERIES** option under the **FORMAT** menu, and then when the Format Data Series dialog box appears, select the **Patterns tab**. Change the **Marker** selection to **None** which erases the data markers, but the trendline remains.

Once the curve for the model is plotted, add the original data points for the actual weight of each bass to the chart by activating the chart, and then selecting **NEW DATA** under the **INSERT** command. The new range of data points is automatically added to the chart. Some editing is required to obtain a plot exactly like the one presented here (see Figure 3.2.8).

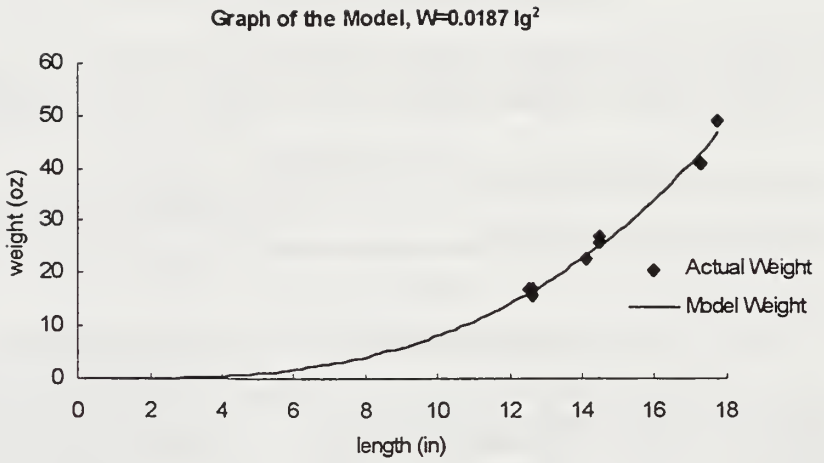


Figure 3.2.8: Graph of the Model, $W = 0.0187\ lg^2$

A fisherman would probably be happier with this new rule, because doubling the girth leads to a fourfold increase in the weight of the fish. However, this model appears more inconvenient to apply. The application of either of the preceding rules would probably require the fisherman to record the length and the girth of each fish, and then compute its weight on a four-function calculator. This method seems awkward for someone in a small boat. Therefore, if a system such as this is to work, the fishing competitors would prefer a device that gave them weight if the length and girth are known. Therefore, no calculations would be required. These model plots would create these devices or charts for the fisherman.

IV. MODEL FITTING AND EXPERIMENTAL MODELING

Chapter III demonstrated Excel's capability to perform various transformations on a data set and to plot the resulting transformed data to assist in determining graphically the adequacy of a proposed proportionality model. In particular, methods to enter data, transform data, obtain a scatterplot, test a proportionality relationship, and estimate the parameters of a model were presented. This chapter describes how to determine the parameters of a model analytically, according to some criterion of "best fit," and to test the adequacy of the model. This procedure is known as *model fitting*. It also introduces a technique to explain the behavior being observed when a model does not exist. In that case, there exists a collection of data points that can be used to predict the behavior within some range of interest by constructing an *experimental* or *empirical model* based solely on the collected data.

When analyzing a collection of data points, there exist two possible cases:

1. Fitting a selected model type or types to the data (using some criterion of "best fit") and then choosing the most appropriate model from competing types that have been fitted.
3. Predicting the behavior within some range of interest using a curve that captures the trend of the data (e.g., one-term, low- or high-order polynomials, or cubic spline models) when no particular type has been identified.

In the model-fitting case, a relationship of a particular type has been identified, and the modeler is willing to accept some deviation between the model and the collected data points in order to have a model that satisfactorily explains the situation under investigation. On the other hand, when interpolating, the modeler is strongly guided by the data that have been carefully collected and analyzed, and a curve is sought that captures the trend of the data to predict in between the data points. In both situations, the modeler usually wants to make predictions from the model. However, the modeler tends to emphasize the proposed models over the data when model fitting, whereas he or she places great confidence in the collected data when interpolating, or experimental modeling (because a particular form of the model is unavailable).

A. MODEL FITTING

Suppose it is proposed that a parabolic model might best explain a behavior being studied. The interest lies in selecting that member of the family $y = Ax^2$ which best fits the given set of data. Using Excel, y versus x^2 can be plotted and a graphical estimate of the slope of the "best" fit line can be determined, as demonstrated in Chapter III. Now we will focus on an analytical method to arrive at an accurate model for a given data set. Again, from the family $y = Ax^2$, 'A' can be determined analytically by using a curve-fitting criterion, such as least-squares or Chebyshev, and solving the resulting optimization problem (see Chapter 5, Giordano and Weir,

op. cit., page 136, for a discussion of model fitting). This section demonstrates how Excel can be used to solve the least-squares optimization problem with analysis of the "goodness of fit" of the resulting model.

1. Least-Squares Curve Fitting Using Excel

The method of least-squares curve fitting is simply the solution to a model such that the sum of the squares of the deviations between the observations and predictions is minimized. Given some function type $y = f(x)$ and a collection of m data points (x_i, y_i), the problem is to determine the parameters of the function type $y = f(x)$ to minimize the sum:

$$\sum_{i=1}^{m} | y_i - f(x_i) |^2$$

In Excel, the method of least-squares curve fitting is performed as linear regression analysis under the **Regression** tool. To locate the **Regression** tool, on the menu bar under the **TOOLS** command is the menu selection **Data Analysis**. By scrolling down and selecting the **Data Analysis** tool, a separate window entitled the Data Analysis dialog box will appear with a list of analysis tools. Using the scroll bar, move down until the selection **Regression** appears (the list appears in alphabetical order). Once the line **Regression** has been highlighted, select the **OK** command to exit the window which immediately enters the **Regression** tool as the **Data Analysis** menu selection. A Regression dialog box will appear which asks for input as well as for what output needs to be generated. The Regression tool performs linear regression analysis.

Regression fits a line through a set of observations using the least squares method. Regression is used in a wide variety of applications that seek to analyze how a single dependent variable is affected by the values of one or more independent variables. In general, regression arrives at an equation for performance based on each of the inputs. Depending on the options chosen when using the **Regression** tool, Microsoft Excel generates the following output:

1. A summary output table.
2. A residuals output table that can include residuals, standardized residuals, and predicted values.
3. A residual plot for each independent variable versus the residual.
4. A line fit plot of the predicted values with the observed values.
5. A normal probability plot.
6. A two-column probability data output table displaying the dependent variable values and percentiles used to generate the normal probability plot.

Be sure to arrange output ranges side-by-side on the worksheet. In general, output tables vary in length rather than width. Allow at least four columns for the residuals output table, and allow at least two columns for the probability data output table.

Example 4.1(a): Vehicular Stopping Distance Revisited

Scenario: Let us reconsider the problem of predicting a motor vehicle's stopping distance as a function of its speed. From Example 3.1, the submodel in which reaction distance d_r was proportional to the velocity v was tested graphically and the constant of proportionality was estimated to be 1.1. Similarly, the submodel predicting a proportionality between braking distance d_b and the square of the velocity was tested. Since there was reasonable agreement with the submodel, it was estimated that the proportionality constant be 0.054. Hence, the model for stopping distance was given by:

$$d = 1.1v + 0.054\, v^2$$

Now fit these two submodels analytically using the Excel Regression command and compare the various fits. Recall the original data given below in Table 4.1.1.

Speed (mph)	Driver Reaction Distance (ft)	Braking Distance (ft)	Average Braking Distance (ft)	Observed Total Stopping Distance (ft)	Average Stopping Distance (ft)
20	22	18	20	40-44	42
25	28	25	28	53-59	56
30	33	36	40.5	69-78	73.5
35	39	47	52.5	86-97	91.5
40	44	64	72	108-124	116
45	50	82	92.5	132-153	142.5
50	55	105	118	160-186	173
55	61	132	148.5	193-226	209.5
60	66	162	182	228-268	248
65	72	196	220.5	268-317	292.5
70	77	237	266	314-372	343
75	83	283	318	366-436	401
80	88	334	376	422-506	464

Table 4.1.1: Observed Reaction and Braking Distance

Using Excel: For Excel to fit the first submodel, $d_r = Av$ (driver reaction distance), the following data is needed from Table 4.1.1: speed and driver reaction distance. Therefore, input this data into Columns A and B of Excel. The data should appear similar to Figure 4.1.1. Excel approximates a curve that captures the trend of the data through regression analysis. Excel arrives at an equation based on each of the inputs.

Now go to the **TOOLS** command on the menu bar and click on the **Data Analysis** command. Once the Data Analysis dialog box appears, scroll down and highlight the **Regression** command, and then press **OK** to enter. The Regression dialog box now appears. Under the **Input Y Range**, ensure the cursor is in the designated box, then go to and highlight Cells B4 through

B16, and then either press the **Tab** button or move the cursor to another box to enter the data. Remember, a dotted line borders the cells to be entered into the range. Also, the **Enter** key is only to be used when all appropriate information has been recorded in the Regression dialog box.

	P	Q
1		Driver Reaction
2	Speed	Distance
3	(mph)	(ft)
4	20	22
5	25	28
6	30	33
7	35	39
8	40	44
9	45	50
10	50	55
11	55	61
12	60	66
13	65	72
14	70	77
15	75	83
16	80	88

Figure 4.1.1: Driver Reaction Distance Data for the Excel Regression Tool

The **Input Y Range** is where the reference for the range of dependent data to be analyzed is entered. The *dependent* data should be typed in a single column and cannot contain more than one piece of data.

Now, move the cursor down to the **Input X Range**, then go to and highlight Cells A4 through A16, and then enter the data in the appropriate manner. The **Input X Range** is where the reference for the range of *independent* data to be analyzed is entered. Excel orders independent variables in ascending order from left to right, using 1,2,3, and so on for the variable names in the summary output table. The maximum number of input ranges for the independent or x variable is 16.

Under the Output options, ensure the **New Worksheet Ply** is marked which places the regression data and plots on a clean worksheet. Under Residuals, mark **Residuals** which includes residuals in the residuals output table (see Figure 4.1.2), mark **Residual Plots** which generates a chart for each independent variable versus the residual, and mark **Line Fit Plots** which generates a chart for predicted values versus the observed values. Finally, mark the **Normal Probability Plot** mainly to obtain the observed data in a probability output table (see Figure 4.1.2). A chart is also generated plotting normal probability plots. Be sure to mark **Constant is Zero** under the Input options since the curve passes through the origin. Once complete, press the **Enter** key or select the **OK** button to execute the regression analysis.

	A	B	C	D	E	F	G
21							
22	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
23							
24	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
25	1	22.08097166	-0.08097166		3.846153846	22	
26	2	27.60121457	0.398785425		11.53846154	28	
27	3	33.12145749	-0.12145749		19.23076923	33	
28	4	38.6417004	0.358299595		26.92307692	39	
29	5	44.16194332	-0.16194332		34.61538462	44	
30	6	49.68218623	0.317813765		42.30769231	50	
31	7	55.20242915	-0.20242915		50	55	
32	8	60.72267206	0.277327935		57.69230769	61	
33	9	66.24291498	-0.24291498		65.38461538	66	
34	10	71.76315789	0.236842105		73.07692308	72	
35	11	77.28340081	-0.28340081		80.76923077	77	
36	12	82.80364372	0.196356275		88.46153846	83	
37	13	88.32388664	-0.32388664		96.15384615	88	
38							

Figure 4.1.2: Residual Output Table (left); Probability Output Table (right)

The following solution to this system is given as coefficients as part of the summary output table located in Figure 4.1.3. The coefficients are as follows

$$a_0 = 0$$

$$a_1 = 1.104048583$$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	0	#N/A
18	X Variable 1	1.104048583	0.001417004
19			

Figure 4.1.3: Partial Summary Output Table

Notice that the deviations or residuals are not very large and are positive as well as negative (see Figure 4.1.4). However, a pattern appears in the residuals plot indicating that the model does not totally explain the behavior. This may indicate that an assumption should have been included in the submodel. The submodel does nevertheless capture the trend of the data (see Figure 4.1.5). (Note: These plots are created automatically by Excel through the Regression tool.)

Originally, based on the principles of physical science, the driver reaction distance was proportional to the velocity. Although the residual plot indicates that something in the model was not accounted for, the scatterplot of the model reveals a straight line through the origin which supports the original claim that driver reaction distance is proportional to the velocity.

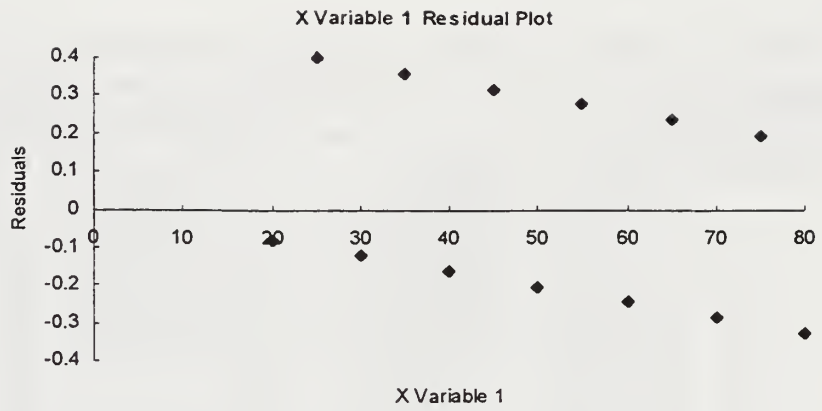


Figure 4.1.4: Plot of the Residuals for the Driver Reaction Distance

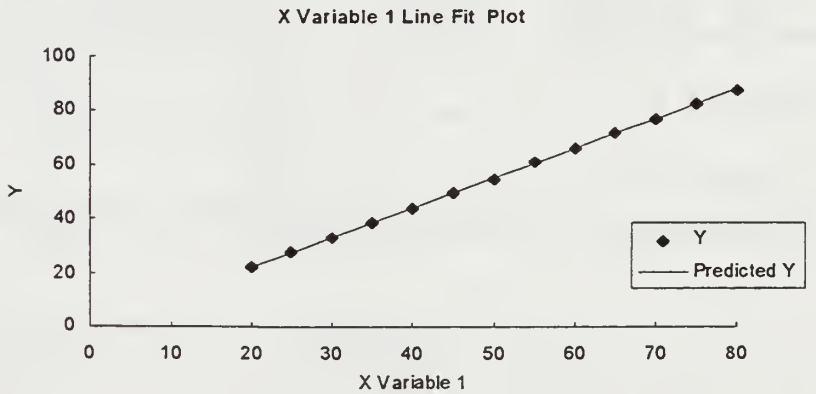


Figure 4.1.5: Plot of the Submodel for the Driver Reaction Distance

Using Excel: (*continued*) For Excel to fit the second submodel, $d_b = Bv^2$ (average braking distance), the following data is needed from Table 4.1.1: speed and average braking distance. Input the average braking distance data into Column D of Excel alongside the data for the first submodel. However, speed squared is needed instead of speed, therefore, use the following Excel formula to create the column for Speed^2 (Column C),

=SUMSQ(A4)

The data should appear similar to Figure 4.1.6.

Now go to the **TOOLS** command on the menu bar and click on the **Data Analysis** command. Once the Data Analysis dialog box appears, scroll down and highlight the **Regression** command, and then press **OK** to enter. The Regression dialog box now appears. Under the **Input Y Range**, ensure the cursor is in the designated box, then go to and highlight Cells D4 through D16, and then either press the **Tab** button or move the cursor to another box to enter the data. Now, move the cursor down to the **Input X Range**, then go to and highlight Cells C4 through C16, and then enter the data in the appropriate manner.

	A	B	C	D
1		Driver Reaction		Average
2	Speed	Distance	Speed^2	Braking
3	(mph)	(ft)	(ft^2)	Distance (ft)
4	20	22	400	20
5	25	28	625	28
6	30	33	900	40.5
7	35	39	1225	52.5
8	40	44	1600	72
9	45	50	2025	92.5
10	50	55	2500	118
11	55	61	3025	148.5
12	60	66	3600	182
13	65	72	4225	220.5
14	70	77	4900	266
15	75	83	5625	318
16	80	88	6400	376
17				

Figure 4.1.6: Average Braking Distance Data for the Excel Regression Tool

Under the Output options, ensure the **New Worksheet Ply** is marked which places the regression data and plots on a clean worksheet. Under Residuals, mark **Residuals** which includes residuals in the residuals output table (see Figure 4.1.7), mark **Residual Plots** which generates a chart for each independent variable versus the residual, and mark **Line Fit Plots** which generates a chart for predicted values versus the observed values. Finally, mark the **Normal Probability Plot** mainly to obtain the observed data in a probability output table (see Figure 4.1.7.).

	A	B	C	D	E	F	G
21							
22	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
23							
24	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
25	1	21.68346256	-1.683462564		3.846153846	20	
26	2	33.88041026	-5.880410256		11.53846154	28	
27	3	48.78779077	-8.287790769		19.23076923	40.5	
28	4	66.4056041	-13.9056041		26.92307692	52.5	
29	5	86.73385026	-14.73385026		34.61538462	72	
30	6	109.7725292	-17.27252923		42.30769231	92.5	
31	7	135.521641	-17.52164103		50	118	
32	8	163.9811856	-15.48118564		57.69230769	148.5	
33	9	195.1511631	-13.15116308		65.38461538	182	
34	10	229.0315733	-8.531573333		73.07692308	220.5	
35	11	265.6224164	0.37758359		80.76923077	266	
36	12	304.9236923	13.07630769		88.46153846	318	
37	13	346.935401	29.06459897		96.15384615	376	
38							

Figure 4.1.7: Residual Output Table (left); Probability Output Table (right)

A chart is also generated plotting normal probability plots. Be sure to mark **Constant is Zero**

under the Input options since the curve passes through the origin. Once complete, press the **Enter** key or select the **OK** button to execute the regression analysis.

The following solution to this system is given as coefficients as part of the summary output table located in Figure 4.1.8. The coefficients are as follows

$$a_0 = 0$$

$$a_1 = 0.054208656$$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	0	#N/A
18	X Variable 1	0.054208656	0.001197877
19			

Figure 4.1.8: Partial Summary Output Table

Notice that the plot of the deviations or residuals appear to be in a shape of a curve. This indicates a pattern, which means that something is not being accounted for in the model (see Figure 4.1.9). However, the residuals look reasonable, and the plot of average braking distance versus speed squared forms a straight line through the origin. Therefore, the submodel does nevertheless capture the trend of the data (see Figure 4.1.10). (Note: These plots are created automatically by Excel through the Regression tool.)

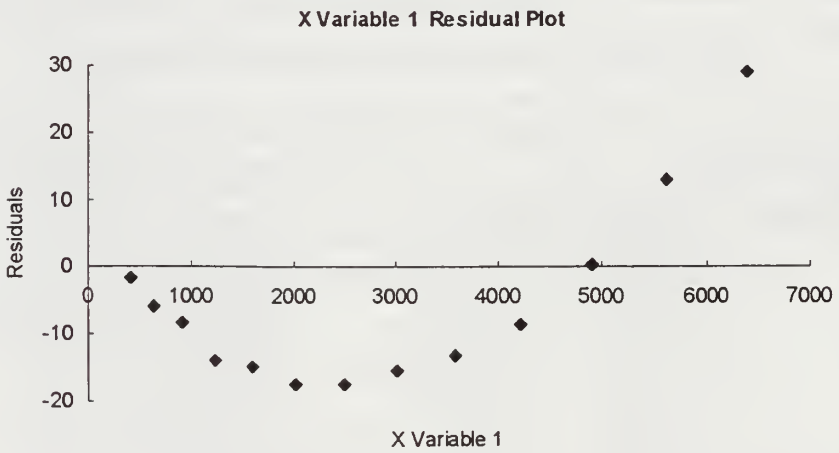


Figure 4.1.9: Plot of the Residuals for the Driver Reaction Distance

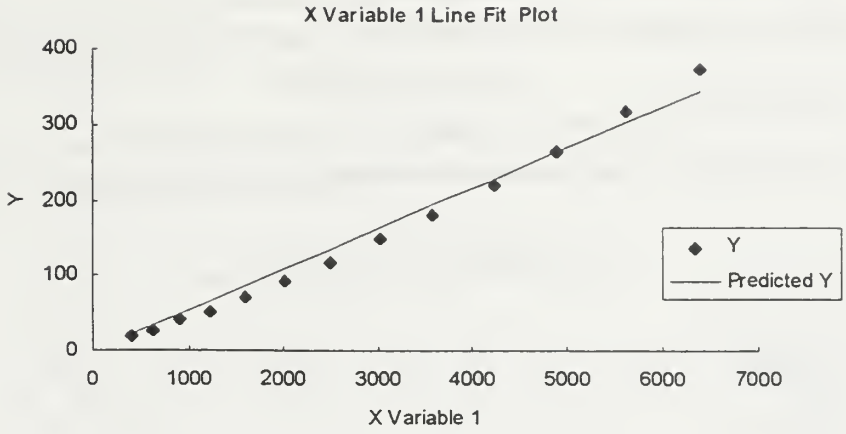


Figure 4.1.10: Plot of the Submodel for the Driver Reaction Distance

Scenario: (*continued*) Considering the grossness of the assumptions and the inaccuracies of the data, the coefficients are rounded to obtain the model (e.g., the two submodels are added together):

$$d = 1.104\, v + 0.0542\, v^2$$

Therefore, the above model does not differ significantly from that obtained graphically in Example 3.1. In order to compare the observed data with the model created from the two submodels, plot the observed data points, which are the average stopping distances, versus the least-squares model for the total stopping distances.

Using Excel: In order to plot this new model (created from adding the two submodels) along with the actual, observed data points, then Column C needs to be created and Column D added first, as depicted in Figure 4.1.11.

	B	C	D	E
1			Predicted (LS)	Observed
2	Speed	Speed^2	Stopping	Average Stopping
3	(mph)	(ft^2)	Distance (ft)	Distance (ft)
4	20	400	43.76	42
5	25	625	61.475	56
6	30	900	81.9	73.5
7	35	1225	105.035	91.5
8	40	1600	130.88	116
9	45	2025	159.435	142.5
10	50	2500	190.7	173
11	55	3025	224.675	209.5
12	60	3600	261.36	248
13	65	4225	300.755	292.5
14	70	4900	342.86	343
15	75	5625	387.675	401
16	80	6400	435.2	464

Figure 4.1.11: Data Computed for the Graphical and Least-Squares Models

Column C is the least-squares model depicted by the equation $d = 1.104\,v + 0.0542\,v^2$, which was obtained by adding the two submodels created just before, and uses the following Excel formula:

=1.104*A4+0.0542*B4

Now, plot the observed data points, which are the average stopping distances (see Figure 4.1.11 for the complete data), versus the least-squares model for the stopping distances (see Figure 4.1.12).

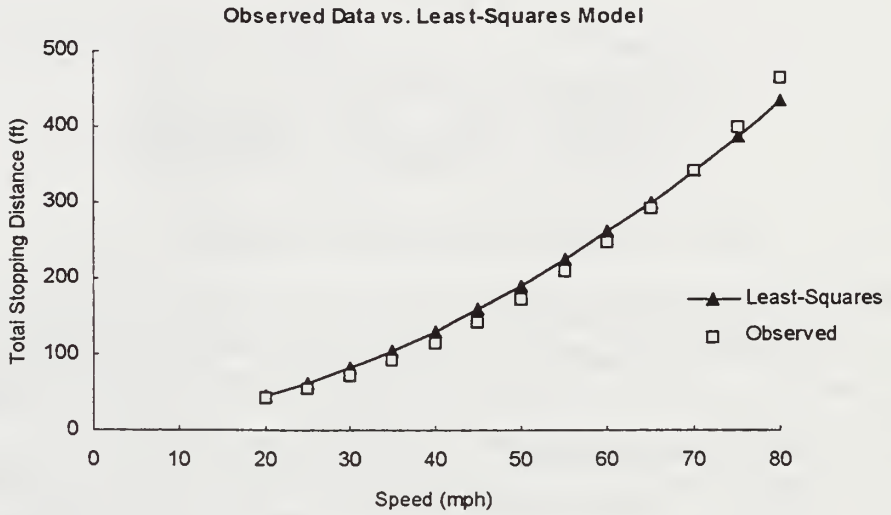


Figure 4.1.12: A Plot of the Least-Squares Model versus the Observed Data Points

Notice how the model, $d = 1.104\,v + 0.0542\,v^2$, captures the trend of the data in Figure 4.1.12.

Now, to find the deviations or residuals of the observed data points from the least-squares model, subtract the observed stopping distance (average) from the model's calculated stopping distances (e.g., Predicted − Actual). The results are found in Figure 4.1.13. Then plot these deviations or residuals versus the speed on a separate chart (see Figure 4.1.14)

	A	B	C	D	E
1			Predicted (LS)	Observed	Predicted (LS)
2	Speed	Speed	Stopping	Average Stopping	Model
3	(mph)	(mph)	Distance (ft)	Distance (ft)	Residuals
4	400	20	43.76	42	1.76
5	625	25	61.475	56	5.475
6	900	30	81.9	73.5	8.4
7	1225	35	105.035	91.5	13.535
8	1600	40	130.88	116	14.88
9	2025	45	159.435	142.5	16.935
10	2500	50	190.7	173	17.7
11	3025	55	224.675	209.5	15.175
12	3600	60	261.36	248	13.36
13	4225	65	300.755	292.5	8.255
14	4900	70	342.86	343	-0.14
15	5625	75	387.675	401	-13.325
16	6400	80	435.2	464	-28.8

Figure 4.1.13: Deviations (or Residuals) from the Observed Data Points and the Model

Now plot the residuals for the least-squares model, located in Column E above, versus the speed (see Figure 4.1.14).

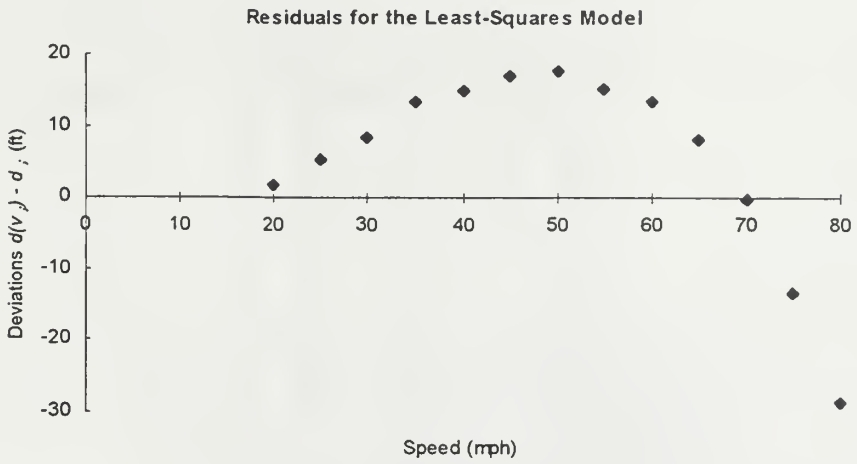


Figure 4.1.14: Plot of the Deviations (or Residuals) for the Least-Squares Model

2. Choosing a Best Model

Another model is obtained by directly fitting the data for total stopping distance to the quadratic

$$d = k_1 v + k_2 v^2$$

instead of fitting each submodel individually. Now use the Regression tool from Excel to model this single equation.

Example 4.1(b): Vehicular Stopping Distance Revisited (continued)

Scenario: (*continued*) For this part, Excel is the only source used to calculate the model as well as plot the model and plot the residuals. Using the same data, the model is represented by a quadratic with thirteen (13) known data points.

Using Excel: Ensure each set of recorded data for the quadratic is entered. The following data is needed: speed, speed^2, and average stopping distance. The data should appear similar to Figure 4.1.15. Now go to the **TOOLS** command on the menu bar and click on the **Data Analysis** command. Once the Data Analysis dialog box appears, scroll down and highlight the **Regression** command, and then press **OK** to enter. The Regression dialog box now appears. Under the **Input Y Range**, ensure the cursor is in the designated box, then go to and highlight Cells C4 through C16, and then either press the **Tab** button or move the cursor to another box to enter the data. Remember, a

dotted line borders the cells to be entered into the range. Now, move the cursor down to the **Input X Range**, then go to and highlight Cells A4 through B16, and then enter the data in the appropriate manner.

	A	B	C
1			Average
2	Speed	Speed^2	Stopping
3	(mph)	(mph)^2	Distance (ft)
4	20	400	42
5	25	625	56
6	30	900	73.5
7	35	1225	91.5
8	40	1600	116
9	45	2025	142.5
10	50	2500	173
11	55	3025	209.5
12	60	3600	248
13	65	4225	292.5
14	70	4900	343
15	75	5625	401
16	80	6400	464

Figure 4.1.15: Vehicle Stopping Distance Data for the Excel Regression Tool

Under the Output options, ensure the **New Worksheet Ply** is marked which places the regression data and plots on a clean worksheet. Under Residuals, mark **Residuals** which includes residuals in the residuals output table (see Figure 4.1.16), mark **Residual Plots** which generates a chart for each independent variable versus the residual, and mark **Line Fit Plots** which generates a chart for predicted values versus the observed values. Finally, mark the **Normal Probability Plot** mainly to obtain the observed data in a probability output table (see Figure 4.1.16).

	A	B	C	D	E	F	G
23	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
24							
25	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
26	1	30.43319317	11.56680683		3.846153846	42	
27	2	46.62691355	9.373086451		11.53846154	56	
28	3	66.25480277	7.245197231		19.23076923	73.5	
29	4	89.31686083	2.183139173		26.92307692	91.5	
30	5	115.8130877	0.186912278		34.61538462	116	
31	6	145.7434835	-3.243483453		42.30769231	142.5	
32	7	179.108048	-6.108048021		50	173	
33	8	215.9067814	-6.406781427		57.69230769	209.5	
34	9	256.1396837	-8.139683669		65.38461538	248	
35	10	299.8067547	-7.306754748		73.07692308	292.5	
36	11	346.9079947	-3.907994664		80.76923077	343	
37	12	397.4434034	3.556596583		88.46153846	401	
38	13	451.412981	12.58701899		96.15384615	464	
39							

Figure 4.1.16: Residual Output Table (left); Probability Output Table (right)

A chart is also generated plotting normal probability plots. Be sure to mark **Constant is Zero** under the Input options since the curve passes through the origin. Once complete, press the **Enter** key or select the **OK** button to execute the regression analysis.

The following solution to this system is given as coefficients as part of the summary output table located in Figure 4.1.17. The coefficients are as follows

$$a_0 = 0$$

$$a_1 = 0.147992123$$

$$a_2 = 0.068683377$$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	0	#N/A
18	X Variable 1	0.147992123	0.170828455
19	X Variable 2	0.068683377	0.002664048
20			
21			

Figure 4.1.17: Partial Summary Output Table

Notice that the deviations or residuals are positive as well as negative (see Figure 4.1.18). Also, a pattern appears in the residuals plot indicating that the model does not totally explain the behavior. The residual plot appears upside down only because Excel subtracts the Predicted Values from the Actual Values, instead of the other way depicted in Figure 4.1.14. However, the model does nevertheless capture the trend of the data even more accurately than the model created by the two submodels (see Figure 4.1.19). (Note: These plots are created automatically by Excel through the Regression tool.)

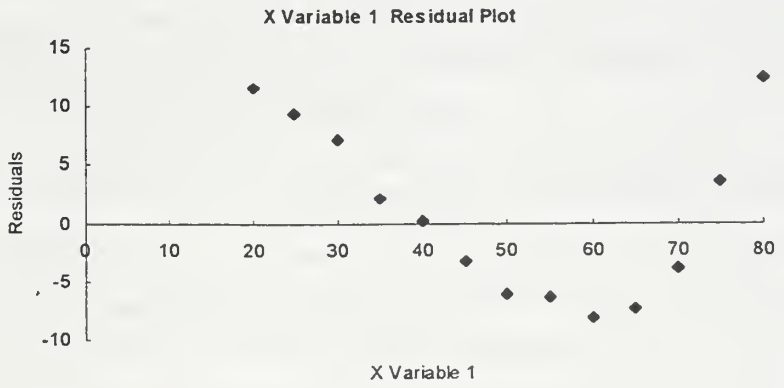


Figure 4.1.18: Plot of the Residuals for the Vehicle Stopping Distance Model

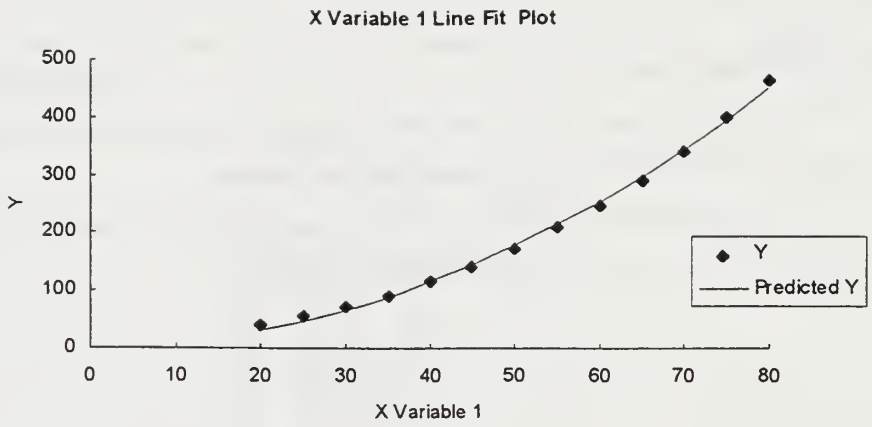


Figure 4.1.19: Plot of the Model for the Vehicle Stopping Distance

B. EXPERIMENTAL MODELING

In previous sections, when fitting a curve, the modeler is using assumptions and, possibly, physical or mathematical principles, to construct a particular type of model which helps explain the behavior being observed. If collected data then justify the soundness of those assumptions, the modeler's next task is to determine the parameters of the selected curve that best fits the data according to some criterion (such as least squares). This was the methodology used for the total stopping distance model, for instance. In this situation the modeler expects, and willingly accepts, some deviations between the fitted model and the collected data to obtain a model explaining the behavior. The problem with this approach is that in many cases the modeler is unable to construct a tractable model form that satisfactorily explains the behavior. Thus the modeler does not know what kind of curve actually describes the behavior.

If it is necessary to predict the behavior, nevertheless, the modeler may conduct experiments (or otherwise gather data) to investigate the behavior of dependent variable(s) for selected values of the independent variable(s) within some range. In essence, the modeler desires to find an *empirical model* based on the *collected* data rather than to select (or construct) a model based on certain assumptions. In such cases the modeler is strongly influenced by the data that have been carefully collected and analyzed, so he or she seeks a curve that captures the trend of the data to *predict* in between the data points. (Extrapolations beyond the range of the collected data may lead to erroneous predictions and conclusions.)

The construction of empirical models is now addressed as well as the selection process for simple one-term models that capture the trend of the data. The smoothing of data is investigated using low-order polynomials, and finally, the technique of cubic spline interpolation is presented, where a distinct cubic polynomial is used across successive pairs of data points to form a single smooth curve. A scatterplot of the data is used first to determine if a trend exists. If a trend in the data is discernible, then the modeler begins with the simplest technique

available and gradually increases the sophistication until an empirical model is developed that satisfies the requirements of the particular application. For a more detailed discussion of empirical model construction, see Chapter 6 of Giordano, Weir, and Fox, op. cit., page 166.

1. High-Order Polynomial Models

Because of their mathematical simplicity, one-term models are limited in their ability to capture the trend of any collection of data, and therefore, in some cases models with more than one term must be considered. Because polynomials are easy to integrate and to differentiate, they are especially popular to use. To fit a high-order polynomial, a system of equations is solved, and therefore determined by forcing the polynomial to satisfy each data point. Therefore, high-order polynomials do not have residuals because they pass through the data points. However, although a polynomial does pass through all the data points (within tolerances of computer round-off error), there is severe oscillation of the polynomial near each end of the interval. This tendency of high-order polynomials to oscillate severely near the end points of the interval is a serious disadvantage to using them.

Example 4.2: Elapsed Time of a Tape Recorder

Scenario: We collected data relating the counter on a particular tape recorder with its elapsed playing time. Suppose we are unable to build an explicative model of this system but are still interested in predicting what may occur. As an example, let us construct a model to predict the amount of elapsed time of a tape recorder as a function of its counter reading. In this example, Excel is the only source used to calculate the model as well as plot the model and plot the residuals.

Thus, let c_i represent the counter reading (in hundreds), and t_i (sec) the corresponding amount of elapsed time. Consider the following data:

c_i (in 100's)	1	2	3	4	5	6	7	8
t_i (sec)	205	430	677	945	1233	1542	1872	2224

The model is represented by a 7th order polynomial with eight known data points from above. The polynomial is denoted symbolically by:

$$P_7(c) = a_0 + a_1c + a_2c^2 + a_3c^3 + a_4c^4 + a_5c^5 + a_6c^6 + a_7c^7$$

The eight data points require that the constants a_i satisfy the system of linear algebraic equations:

$$205 = a_0 + a_1(1) + a_2(1)^2 + a_3(1)^3 + a_4(1)^4 + a_5(1)^5 + a_6(1)^6 + a_7(1)^7$$

$$430 = a_0 + a_1(2) + a_2(2)^2 + a_3(2)^3 + a_4(2)^4 + a_5(2)^5 + a_6(2)^6 + a_7(2)^7$$

$$\vdots$$

$$2224 = a_0 + a_1(8) + a_2(8)^2 + a_3(8)^3 + a_4(8)^4 + a_5(8)^5 + a_6(8)^6 + a_7(8)^7$$

Using Excel: We can use the Regression command to find the coefficients of the polynomial $P_7(c)$ which passes exactly through all the data points. However, this technique normally cannot be done. A program is needed to solve a system of equations, like Matlab. Excel does not have the capability to solve a system of equations. However, in this case, the approximate solution is the actual solution. Therefore, the Regression tool in Excel can be used. The first step is to enter the data for the polynomial for each set of recorded data. It should appear similar to Figure 4.2.1.

	A	B	C	D	E	F	G	H	I
1									
2	Polynomial	1st Term	2nd Term	3rd Term	4th Term	5th Term	6th Term	7th Term	Y Value
3	1st data point	1	1	1	1	1	1	1	205
4	2nd data point	2	4	8	16	32	64	128	430
5	3rd data point	3	9	27	81	243	729	2187	677
6	4th data point	4	16	64	256	1024	4096	16384	945
7	5th data point	5	25	125	625	3125	15625	78125	1233
8	6th data point	6	36	216	1296	7776	46656	279936	1542
9	7th data point	7	49	343	2401	16807	117649	823543	1872
10	8th data point	8	64	512	4096	32768	262144	2097152	2224
11									

Figure 4.2.1: Elapsed Tape Recorder Data

Now go to the **TOOLS** command on the menu bar and click on the **Data Analysis** command. Once the Data Analysis dialog box appears, scroll down and highlight the **Regression** command, and then press **OK** to enter. The Regression dialog box now appears. Under the **Input Y Range**, ensure the cursor is in the designated box, then go to and highlight Cells I3 through I10 (which are the dependent variables), and then either press the **Tab** button or move the cursor to another box to enter the data. Remember, a dotted line borders the cells to be entered into the range. The dependent data should be typed in a single column and cannot contain more than one piece of data. Now, move the cursor down to the **Input X Range**, then go to and highlight Cells B3 through H10 (which are the independent variables), and then enter the data in the appropriate manner.

Under the Output options, ensure the **New Worksheet Ply** is marked which places the regression data and plots on a clean worksheet. Under Residuals, mark **Residuals** which includes residuals in the residuals output table (see Figure 4.2.2), mark **Residual Plots** which generates a chart for each independent variable versus the residual, and mark **Line Fit Plots** which generates a chart for predicted values versus the observed values. Finally, mark the **Normal Probability Plot** mainly to obtain the observed data in a probability output table (see Figure 4.2.2). A chart is also generated plotting normal probability plots. Once complete, press the **Enter** key or select the **OK** button to execute the regression analysis.

	A	B	C	D	E	F	G
28	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
29							
30	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
31	1	205	-1.31042E-08		6.25	205	
32	2	429.9999999	7.66249E-08		18.75	430	
33	3	677.0000002	-1.94981E-07		31.25	677	
34	4	944.9999997	2.79488E-07		43.75	945	
35	5	1233	-2.43801E-07		56.25	1233	
36	6	1542	1.29297E-07		68.75	1542	
37	7	1872	-3.8541E-08		81.25	1872	
38	8	2224	4.97039E-09		93.75	2224	
39							

Figure 4.2.2: Residual Output Table (left); Probability Output Table (right)

In Figure 4.2.2, notice how the predicted values for y are in agreement with the observed data, a requirement since high-order polynomial passes exactly through the data points. The following solution to this system is given as coefficients as part of the summary output table located in Figure 4.2.3. The coefficients are as follows:

$a_0 = -13.9999923$ $\quad a_4 = -5.354166491$

$a_1 = 232.9119031$ $\quad a_5 = 0.8013888621$

$a_2 = -29.08333188$ $\quad a_6 = -0.0624999978$

$a_3 = 19.78472156$ $\quad a_7 = 0.0019841269$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	-13.99994968	0
18	X Variable 1	232.9117811	0
19	X Variable 2	-29.08321775	0
20	X Variable 3	19.78466766	0
21	X Variable 4	-5.354152374	0
22	X Variable 5	0.801386784	0
23	X Variable 6	-0.062499837	0
24	X Variable 7	0.001984122	0
25			

Figure 4.2.3: Partial Summary Output Table

Also notice that the deviations or residuals are very small and are positive as well as negative (see Figure 4.2.4). In fact, since the polynomial passes through all the data points, all residuals should be zero (0), except for computer round-off error. For the first time, the residual plot does not have a set pattern. This indicates that possibly all assumptions were correct, and therefore, the behavior of the model is quite accurate. This model captures the trend of the data as depicted in Figure 4.2.5. (Note: These plots are created automatically by the Excel Regression tool.)

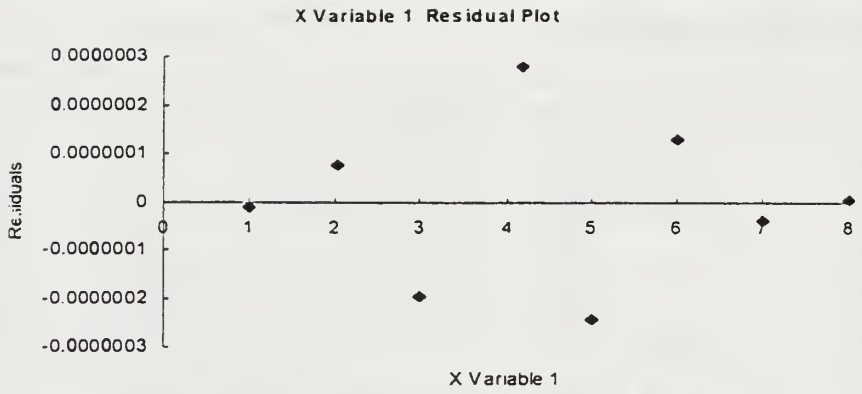


Figure 4.2.4: Plot of the Residuals for the Elapsed Time of a Tape Recorder

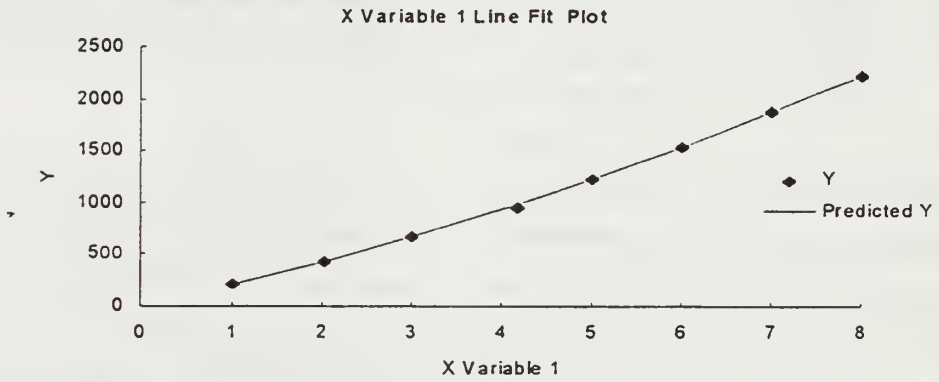


Figure 4.2.5: Plot of the Model for the Elapsed Time of a Tape Recorder

NOTE: The Regression polynomial, in this case, happens to be the actual polynomial when it is solved as a system of equations. Least-squares regression is a polynomial itself. This can be used as long as the polynomial is of degree ≤ 14.

2. Smoothing: Low-Order Polynomial Models

Smoothing with low-order polynomials is an attempt to retain the advantages of polynomials as empirical models while at the same time reducing the tendencies of higher-order polynomials to snake and oscillate. This choice normally results in a situation in which the number of data points exceeds the number of constants necessary to determine the best-fitting polynomial. Because there are fewer constants to determine than there are data points, the low-order polynomial generally does not pass through all the data points. The combination of using a low-order polynomial, while not requiring that it pass through each data point, reduces both the tendency of the polynomial to oscillate and its sensitivity to small changes in the data. When considering the use of a low-

order polynomial for smoothing, two issues come to mind:

1. Should a polynomial be used?
2. If so, what order of polynomial would be appropriate?

Therefore, the derivative concept can help in answering these two questions.

a. Divided Differences

An *nth* order polynomial is characterized by the properties that its *nth* derivative is constant and its *(n+1)st* derivative is identically zero. Nevertheless, if dy/dx is to be zero, then Δy must go to zero. Thus, the differences $y_{i+1} - y_i = \Delta y$ can be computed between successive function values in a designated tabled data to gain insight into what the first derivative is doing. Likewise, because the first derivative is itself a function, the process can be repeated to estimate the second derivative. That is, the differences between successive estimates of the first derivative can be computed to approximate the second derivative. Therefore, the divided difference table is composed of two columns of data and successive columns of divided differences which approximate derivatives. The first divided difference can be interpreted as the difference between two data points divided by the length of the interval over which the change has taken place. Therefore, the denominator for each stage of the divided difference table is the length of change in the interval.

The first differences, denoted by Δ, are constructed by computing $y_{i+1} - y_i$ for $i = 1,2,3,...,n$. The second differences, denoted by Δ^2, are computed by finding the difference between successive first differences from the Δ column. The process can be continued, column by column, until Δ^{n-1} is computed for n data points. In practice, it is easy to construct a divided difference table. The next-higher-order divided difference is generated by taking differences between adjacent current order divided differences and dividing them by the length of the interval over which the change has taken place. The objective here is to use a divided difference table as a qualitative aid to determine if a low-order polynomial is worthy of further investigation.

Example 4.3: Elapsed Time of a Tape Recorder Revisited

Scenario: Returning now to the construction of an empirical model for the elapsed time for a tape recorder, how might the order of the low-order, smoothing polynomial be chosen? Let us begin by constructing the divided difference table for the given data in the table below.

c_i	100	200	300	400	500	600	700	800
t_i (sec)	205	430	677	945	1233	1542	1872	2224

Using Excel: First enter the data for c_i in Column A, and the data for t_i in Column B on a new worksheet. Give Column A a heading of x_i, Column B a heading of y_i, and any other headings that seem appropriate. In Excel, however, the divided difference table will not have the pyramid look or similar affect due to the fact that Excel does everything in block alignment. Excel can calculate all

the computations, but the appearance of the divided difference table will be triangular. The key Excel formulas when calculating divided differences are the very first formulas in each column. Once these formulas are correct, the remaining cells in a column can be updated and changed by simply using the fill handle to copy the formulas into the other cells.

In Cell C3, enter the following Excel formula:

=(B4-B3)/(A4-A3)

then use the fill handle to copy this formula in Column C down through Cell C9. Column C is now complete. Move to Column D and in Cell D3, enter the following Excel formula:

=(C4-C3)/(A5-A3)

Again, use the fill handle to copy this formula in Column D down through Cell D8. In Cell E3, enter the following Excel formula:

=(D4-D3)/(A6-A3)

After copying this formula in Column E down through Cell E7, move to Column F and in Cell F3, enter the following Excel formula:

=(E4-E3)/(A7-A3)

Finally, copy this formula in Column F down through Cell F6, and the divided difference table for the elapsed time of a tape recorder is complete (see Figure 4.3.1).

	A	B	C	D	E	F
1	Data		Divided Differences			
2	x_i	y_i	(delta)	$(delta)^2$	$(delta)^3$	$(delta)^4$
3	100	205	2.2500	0.0011	0.0000	0.0000
4	200	430	2.4700	0.0011	0.0000	0.0000
5	300	677	2.6800	0.0010	0.0000	0.0000
6	400	945	2.8800	0.0011	0.0000	0.0000
7	500	1233	3.0900	0.0011	0.0000	
8	600	1542	3.3000	0.0011		
9	700	1872	3.5200			
10	800	2224				

Figure 4.3.1: Divided Difference Table for the Elapsed Time of a Tape Recorder

Notice that the third column of the divided difference table is all zeros (0's). Therefore, drawing a conclusion from this observation, let us try to fit a quadratic equation to the tape recorder data since the third derivative of a quadratic is zero (0). The quadratic model is of the following form:

$$P_2(c) = a + bc + dc^2$$

where c is the counter reading, $P_2(c)$ is the elapsed time, and a, b, and d are constants to be determined.

Using the data from Columns A and B in Figure 4.3.1, insert an additional column that squares the counter reading, and then apply the **Regression** tool from Excel to this data (see Figure 4.3.2).

	A	B	C
1			Elapsed Time
2	c_1	c_1^2	$P_2(c)$
3	100	10000	205
4	200	40000	430
5	300	90000	677
6	400	160000	945
7	500	250000	1233
8	600	360000	1542
9	700	490000	1872
10	800	640000	2224

Figure 4.3.2: Tape Recorder Data for Quadratic Equation

The residuals are in Figure 4.3.4 and the following solution for the coefficients is given as part of the summary output table located in Figure 4.3.3. The coefficients are as follows:

$$a = 0.1442857143$$

$$b = 1.942261905$$

$$d = 0.001046429$$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	0.142857143	0.593473575
18	X Variable 1	1.942261905	0.003025777
19	X Variable 2	0.001046429	3.28192E-06
20			

Figure 4.3.3: Partial Summary Output Table

	A	B	C	D	E	F	G
22							
23	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
24							
25	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
26	1	204.8333333	0.166666667		6.25	205	
27	2	430.452381	-0.452380952		18.75	430	
28	3	677	-7.95808E-13		31.25	677	
29	4	944.4761905	0.523809524		43.75	945	
30	5	1232.880952	0.119047619		56.25	1233	
31	6	1542.214286	-0.214285714		68.75	1542	
32	7	1872.47619	-0.476190476		81.25	1872	
33	8	2223.666667	0.333333333		93.75	2224	

Figure 4.3.4: Residual Output Table (left); Probability Output Table (right)

The residual plot and line fit plot are in Figures 4.3.5 and 4.3.6, respectively. These plots are

automatically created by the Excel regression tool. Only editing is required in order to make the plots more presentable.

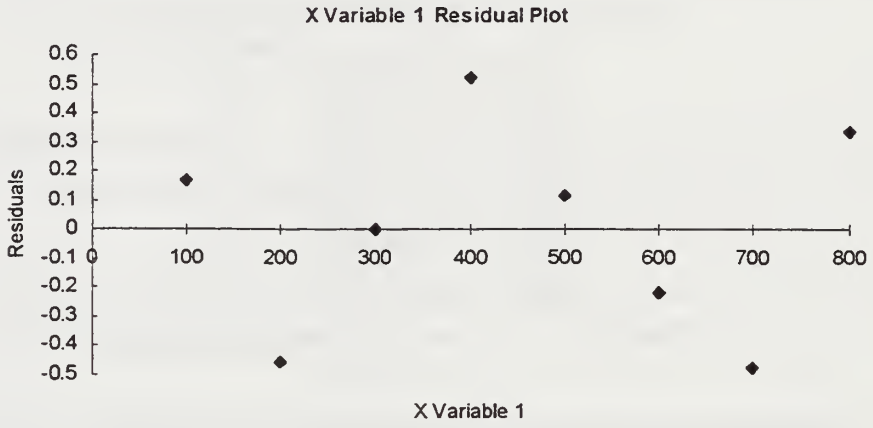


Figure 4.3.5: Plot of the Residuals for the Elapsed Time of a Tape Recorder

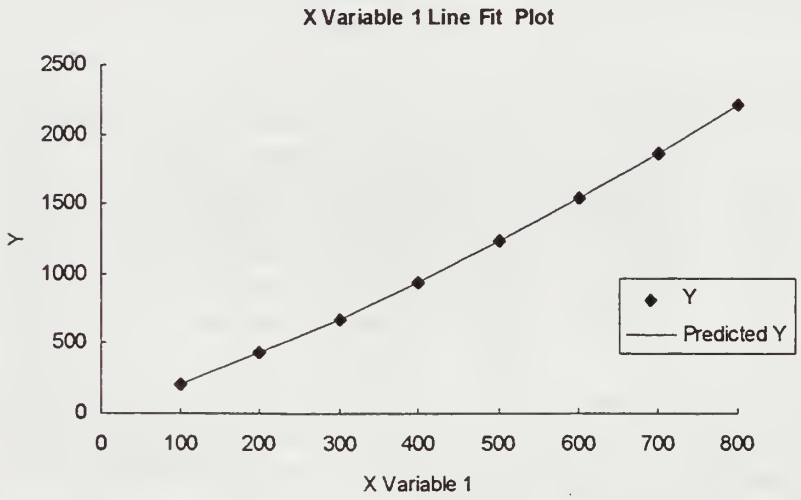


Figure 4.3.6: Plot of the Model versus the Observed Data for the Elapsed Time of a Tape Recorder

The residuals in this case are quite small and are not significant enough to change the behavior of the model. From Figure 4.3.6, the quadratic seems to be a good fit for the model using the following equation:

$$P_2(c) = 0.1442857143 + 1.942261905\,c + 0.001046429\,c^2$$

b. Observations on Difference Tables

Several observations about divided difference tables are in order. First, the x_i's must be distinct and listed in increasing order. It is important to be sensitive to x_i's that are close together because division by a small number can cause numerical difficulties. The scales used to measure both the x_i and y_i must also be considered. For example, suppose the x_i represent distances and are currently measured in miles. If the units are changed to feet, the denominators become larger, resulting in divided differences that are much smaller. Thus judgement on what is small is relative and qualitative.

Example 4.4: Growth of a Yeast Culture Revisited

Scenario: In this example, consider a collection of data points for which a divided difference table can help in deciding whether a low-order polynomial will provide a satisfactory empirical model. The data represent the population of yeast cells in a culture measured over time (in hours). Construct a divided difference table for the data given in Table 4.4.1.

	Data
t_i (hrs)	P_i
0	9.60
1	18.30
2	29.00
3	47.20
4	71.10
5	119.10
6	174.60
7	257.30
8	350.70
9	441.00
10	513.30
11	559.70
12	594.80
13	629.40
14	640.80
15	651.10
16	655.90
17	659.60
18	661.80

Table 4.4.1: Population of Yeast Cells in a Culture Measured in Hours

Using Excel: Just like before, first enter the data for t_i in Column A, and the data for P_i in Column B on a new worksheet. Give Column A a heading of t_i which represents time in hours, Column B a heading of P_i which represents the population, and any other headings that seem appropriate. Remember in Excel, the divided difference table will be triangular. (Note: Once a divided difference table has been created in Excel, it can be used over and over again by simply inputting

the data in the same columns as the previous data and then extending the columns with the formulas as necessary. Therefore, no new formulas need to be created, and the data with results are automatically updated. When this new divided difference table is saved, save it under a new file name in order to keep the original file.)

If starting in a new worksheet, then in Cell C3, enter the following Excel formula:

=(B4-B3)/(A4-A3)

then use the fill handle to copy this formula in Column C down through Cell C20. Column C is now complete. Move to Column D and in Cell D3, enter the following Excel formula:

=(C4-C3)/(A5-A3)

Again, use the fill handle to copy this formula in Column D down through Cell D19. In Cell E3, enter the following Excel formula:

=(D4-D3)/(A6-A3)

After copying this formula in Column E down through Cell E18, move to Column F and in Cell F3, enter the following Excel formula:

=(E4-E3)/(A7-A3)

Finally, copy this formula in Column F down through Cell F17, and the divided difference table for the population of yeast cells in a culture is complete (see Figure 4.4.1).

	A	B	C	D	E	F
1	Data		Divided Differences			
2	t_i (hrs)	P_i	(delta)	$(\text{delta})^2$	$(\text{delta})^3$	$(\text{delta})^4$
3	0	9.60	8.70	1.00	0.92	-0.30
4	1	18.30	10.70	3.75	-0.30	0.84
5	2	29.00	18.20	2.85	3.07	-1.46
6	3	47.20	23.90	12.05	-2.77	1.51
7	4	71.10	48.00	3.75	3.28	-1.51
8	5	119.10	55.50	13.60	-2.75	0.11
9	6	174.60	82.70	5.35	-2.30	-0.05
10	7	257.30	93.40	-1.55	-2.48	0.29
11	8	350.70	90.30	-9.00	-1.32	0.94
12	9	441.00	72.30	-12.95	2.43	-0.16
13	10	513.30	46.40	-5.65	1.80	-1.40
14	11	559.70	35.10	-0.25	-3.78	1.87
15	12	594.80	34.60	-11.60	3.68	-1.10
16	13	629.40	11.40	-0.55	-0.73	0.37
17	14	640.80	10.30	-2.75	0.73	-0.20
18	15	651.10	4.80	-0.55	-0.07	
19	16	655.90	3.70	-0.75		
20	17	659.60	2.20			
21	18	661.80				

Figure 4.4.1: Divided Difference Table for the Population of a Yeast Culture

Scenario: (*continued*) Note that the first divided differences Δ are increasing until $t = 8$ hours, when they begin to decrease. Create a scatterplot of the data. The divided difference table in Figure 4.4.1 suggests that a quadratic function would not be a good model (because of the point of inflection suggested by the sign change in Column D, the second divided differences), so we try to fit a cubic polynomial instead. Using the least-squares criterion, let us develop a cubic model for the data in Figure 4.4.1, and then plot the model and the data points on the same chart. We also plot the residuals on a separate chart.

Using Excel: First, the cubic model is of the following form:

$$P_i = a_0 + a_1 t_i + a_2 t_i^2 + a_3 t_i^3$$

Therefore, Column A, which is the data for t_i, needs to be squared and cubed in order for Excel to conduct the linear regression for a cubic correctly. Go to the Column B and Column C headings and highlight them. Then under the **INSERT** command, select **Columns** to insert two new columns. Excel inserts a new Column B and Column C, moves all other columns over to the right by two columns, and names them Column D, Column E, Column F, etc.. Now in Cell B3, enter the following Excel formula which squares Column A:

=(A3)^2

Again, copy this formula from Cell B3 down through Cell B21 using the fill handle. Now in Cell C3, enter the following Excel formula which cubes Column A:

=(A3)^3

The data is now ready to use Excel's regression tool (see Figure 4.4.2).

	A	B	C	D
1		Data	Data	Observed
2	t_i (hrs)	t_i^2 (hr)^2	t_i^3 (hr)^3	P_i
3	0	0	0	9.60
4	1	1	1	18.30
5	2	4	8	29.00
6	3	9	27	47.20
7	4	16	64	71.10
8	5	25	125	119.10
9	6	36	216	174.60
10	7	49	343	257.30
11	8	64	512	350.70
12	9	81	729	441.00
13	10	100	1000	513.30
14	11	121	1331	559.70
15	12	144	1728	594.80
16	13	169	2197	629.40
17	14	196	2744	640.80
18	15	225	3375	651.10
19	16	256	4096	655.90
20	17	289	4913	659.60
21	18	324	5832	661.80

Figure 4.4.2: Yeast Culture Data Computed for Least-Squares Model

To use the Regression tool for the least-squares criterion, go to the **TOOLS** command on the menu bar and click on the **Data Analysis** command. Once the Data Analysis dialog box appears, scroll down and highlight the **Regression** command, and then press **OK** to enter. The Regression dialog box now appears. Under the **Input Y Range**, ensure the cursor is in the designated box, then go to and highlight Cells D3 through D21, and then either press the **Tab** button or move the cursor to another box to enter the data. Remember, a dotted line borders the cells indicating that this data is the range designated for entry. (Note: The **Enter** key is only used when all appropriate information has been recorded in the Regression dialog box.) Now, move the cursor down to the **Input X Range**, then go to and highlight Cells A3 through C21, and then enter the data in the appropriate manner.

Under the Output options, ensure the **New Worksheet Ply** is marked which places the regression data and plots on a clean worksheet. Under Residuals, mark **Residuals** which includes residuals in the residuals output table (see Figure 4.4.3), mark **Residual Plots** which generates a chart for each independent variable versus the residual, and mark **Line Fit Plots** which generates a chart for predicted values versus the observed values. Finally, mark the **Normal Probability Plot** mainly to obtain the observed data in a probability output table (see Figure 4.4.3). A chart is also generated plotting normal probability plots. Once complete, press the **Enter** key or select the **OK** button to execute the regression analysis. In Figure 4.4.3, notice how the predicted values for y are not in agreement with the observed data. The model fits poorly as expected and fails to capture the trend of the data.

	A	B	C	D	E	F	G
23							
24	RESIDUAL OUTPUT				PROBABILITY OUTPUT		
25							
26	*Observation*	*Predicted Y*	*Residuals*		*Percentile*	*Y*	
27	1	8.213328776	1.386671224		2.631578947	9.6	
28	2	4.770608339	13.52939166		7.894736842	18.3	
29	3	19.08821921	9.911780789		13.15789474	29	
30	4	48.66543665	-1.465436653		18.42105263	47.2	
31	5	91.00153593	-19.90153593		23.68421053	71.1	
32	6	143.5957923	-24.49579229		28.94736842	119.1	
33	7	203.947481	-29.347481		34.21052632	174.6	
34	8	269.5558773	-12.25587733		39.47368421	257.3	
35	9	337.9202565	12.77974348		44.73684211	350.7	
36	10	406.5398939	34.46010615		50	441	
37	11	472.9140646	40.38593543		55.26315789	513.3	
38	12	534.5420439	25.15795605		60.52631579	559.7	
39	13	588.9231072	5.876892767		65.78947368	594.8	
40	14	633.5565297	-4.156529693		71.05263158	629.4	
41	15	665.9415866	-25.14158659		76.31578947	640.8	
42	16	683.5775532	-32.47755317		81.57894737	651.1	
43	17	683.9637047	-28.06370472		86.84210526	655.9	
44	18	664.5993165	-4.999316473		92.10526316	659.6	
45	19	622.9836637	38.8163363		97.36842105	661.8	
46							

Figure 4.4.3: Residual Output Table (left); Probability Output Table (right)

The following solution for the coefficients is given as part of the summary output table located in Figure 4.4.4. The coefficients are as follows:

$$a_0 = 8.213328776$$

$$a_1 = -13.15646101$$

$$a_2 = 10.13052802$$

$$a_3 = -0.416787457$$

	A	B	C
15			
16		*Coefficients*	*Standard Error*
17	Intercept	8.213328776	19.65852306
18	X Variable 1	-13.15646101	9.721421238
19	X Variable 2	10.13052802	1.275426495
20	X Variable 3	-0.416787457	0.046521509
21			

Figure 4.4.4: Partial Summary Output Table

Also notice that the deviations or residuals are quite large, both positive as well as negative (see Figure 4.4.5). The particular pattern plotted by the residuals gives an impression that the model does not account for everything, but the line curve versus the observed data points fits quite nicely. This residual plot is automatically created by the Excel regression tool. Only editing is required in order to make the plot more presentable.

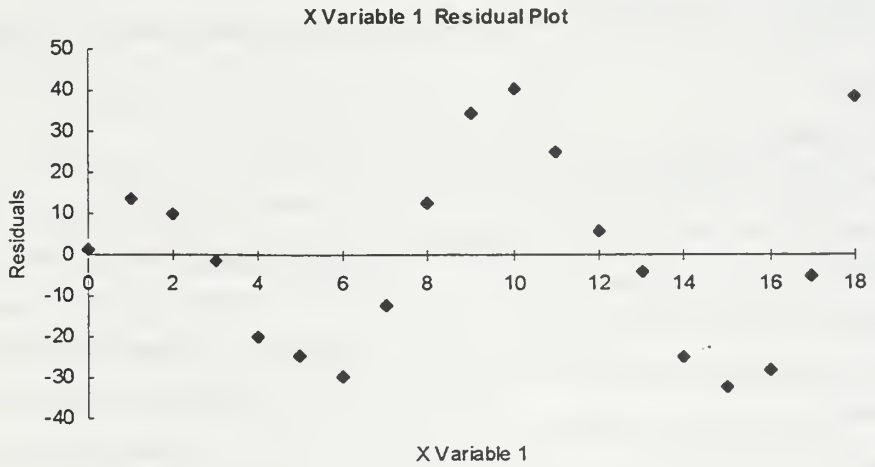


Figure 4.4.5: Plot of the Residuals for the Growth of a Yeast Culture

Finally, to plot the original data with the cubic model on the same chart, the following cubic equation obtained from the linear regression analysis is used:

$$P = 8.213328776 - 13.15646101\,t + 10.13052802\,t^2 - 0.416787457\,t^3$$

Excel automatically creates this plot when the **Line Fit Plots** is marked under *Residuals* in the Regression dialog box. A chart is then generated for predicted values (which are formed from the cubic equation) versus the observed values (which are given, P_i). Again, only editing is required in order to make the plot more presentable (see Figure 4.4.6). For instance, the cubic model is originally plotted as a scatterplot, but to make it easier to see the shape of the curve, highlight the cubic model data points on the chart, and then edit the type of pattern from Marker to Line. It is easy to see that the model does not quite capture the exact trend of the data.

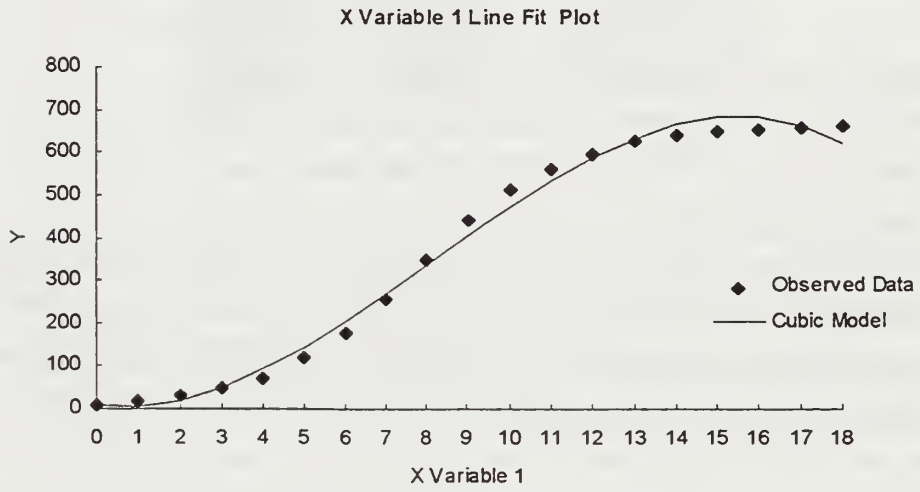


Figure 4.4.6: Plot of the Model versus the Observed Data for the Growth of a Yeast Culture

3. Cubic Spline Models

The use of polynomials in constructing empirical models that capture the trend of the data is appealing because polynomials are so easy to integrate and differentiate. High-order polynomials, however, tend to oscillate near the endpoints of the data interval, and the coefficients can be sensitive to small changes in the data. Unless the data are essentially quadratic or cubic in nature, smoothing with a low-order polynomial may yield a relatively poor fit somewhere over the range of the data. Historically, divided difference tables were used to determine various forms of interpolating polynomials that passed through a chosen subset of data points. Today, other interpolating techniques, such as cubic splines, are more popular. By using different cubic polynomials between successive pairs of data points, the trend of the data can be captured regardless of the nature of the underlying sensitivity to changes in the data.

Cubic splines offer the possibility of matching up not only the slopes but also the curvatures at each interior data point. To determine the constants defining each cubic spline segment, we appeal to the requirement that

each spline pass through the two data points specified by the interval over which the spline is defined. The first and second derivatives of adjacent splines are forced to match at the interior data points, and either the *clamped* or *natural* conditions are applied at the two exterior data points (see Chapter 6.4, Giordano, Weir, and Fox, op. cit., page 201). Unfortunately, Microsoft Excel does not have a built-in cubic spline function. Therefore, we are forced to use an alternative program or tool in order to utilize the cubic spline methodology. The program used is Maple (a computer algebra system). (See Beauchamp, op. cit., page 87.)

Example 4.5: Vehicle Stopping Distance Revisited Again

Scenario: Predict a vehicle's total stopping distance as a function of its speed. The model should have the form:

$$d = k_1 v + k_2 v^2$$

where d is the total stopping distance, v the velocity, and k_1 and k_2 are constants of proportionality resulting from the submodels for reaction distance and mechanical braking distance, respectively. However, since models of this form have been unsatisfactory, yet we are reasonably satisfied with the collected data, let us construct a cubic spline model for the data presented in Table 4.5.1.

Speed, v (mph)	20	25	30	35	40	45	50	55	60	65	70	75	80
Distance, d (ft)	42	56	73.5	91.5	116	142.5	173	209.5	248	292.5	343	401	464

Table 4.5.1: Data Relating Total Stopping Distance and Speed

Using Maple: Since Excel does not have a built-in cubic spline function, we use the following commands in the Microsoft Windows version of Maple. In the Maple system, enter the following commands:

```
> speed:=[20,25,30,35,40,45,50,55,60,65,70,75,80]:
> distance:=[42,56,73.5,91.5,116,142.5,173,209.5,248,292.5,343,401,464]:
> readlib(spline):
> spline(speed, distance, x, cubic):
> s:=unapply(", x):
> sd:={seq([speed[i], distance[i]], i=1..13)}:
> plot1:=plot(s(x), x=20..35):
> plot2:=plot(sd, style=point, symbol=box):
> with(plots):
> display({plot1, plot2});
```

The plot for the first three cubic spline equations should appear similar to Figure 4.5.1.

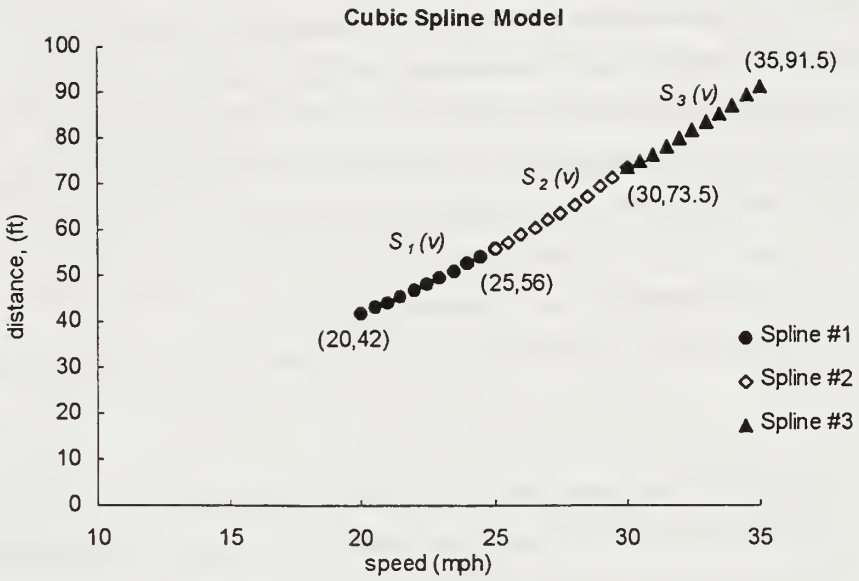


Figure 4.5.1: Plot of the First Three Cubic Spline Equations

V. MONTE CARLO SIMULATION MODELING

A mathematical modeler may encounter situations where the construction of an analytic model is infeasible due to the complexity of the situation. In instances where the behavior cannot be modeled analytically, or data collected directly, the modeler might simulate the behavior indirectly in some manner, and then test various alternatives being considered to estimate how each affects the behavior. Data can then be collected to determine which alternative is best. One type of simulation is Monte Carlo simulation and is typically accomplished with the aid of a computer. In this chapter, Excel is used to construct Monte Carlo simulations of deterministic and probabilistic behaviors. There are a number of serious mathematical concerns associated with the construction and interpretation of Monte Carlo simulations (see Giordano, Weir, and Fox, op. cit., Chapter 7, page 217). Only the Excel implementation of a Monte Carlo simulation is discussed here.

A behavior being modeled can be either *deterministic* or *probabilistic*. Processes with an element of chance involved are probabilistic (as opposed to deterministic) processes. Monte Carlo simulation provides a probabilistic model. For instance, the area under a curve is deterministic (even though the area may be impossible to find it precisely). On the other hand, the time between arrivals of customers at an elevator on a particular morning is probabilistic behavior. A deterministic model can be used to approximate either a deterministic or probabilistic behavior. Similarly, a Monte Carlo simulation can be used to approximate a deterministic behavior (which can be seen with a Monte Carlo approximation to an area under a curve) or a probabilistic one. The real power of Monte Carlo simulation lies in modeling a probabilistic behavior.

A. RANDOM NUMBER GENERATION USING EXCEL

When any Monte Carlo simulation is performed, random numbers are used. Loosely speaking, a sequence of random numbers uniformly distributed in an interval m to n is a set of numbers with no apparent pattern, where each number between m and n can appear with equal likelihood. Excel has a built-in random number generating function. To locate the **Random Number Generation** tool, on the Menu bar under the **TOOLS** command find the menu selection **Data Analysis**. By scrolling down and selecting the **Data Analysis** tool, a separate window entitled the Data Analysis dialog box appears with a list of analysis tools (the same location where the **Regression** command is found in Chapter IV). Using the scroll bar, move down until the selection **Random Number Generation** appears (the list appears in alphabetical order). Once the entry **Random Number Generation** is highlighted, select the **OK** command to exit the window and immediately enter the **Random Number Generation** tool in the **Data Analysis** menu selection. A Random Number Generation dialog box now appears which asks for input and the specificity of the output to be generated.

The **Random Number Generation** tool fills a given range (e.g., designated number of columns and rows) in a worksheet with independent random numbers drawn from one of several distributions (e.g., Uniform. Discrete, Normal, etc.). The Random Number Generation dialog box first asks for the **Number of Variables**, which is the number of values or coordinates to be placed in the same number of columns in the output table by Excel. If the number of variables is not typed, Excel fills all output table columns. Then, it asks for the **Number of Random Numbers**, which is the number of data points or independent random numbers to be generated for each variable. Each data point is placed in a row of the output table. If the number of data points is not entered, Excel fills all output table rows.

Excel then asks for the type of **Distribution** to use to create the random variables. The choices are as follows: **Uniform, Normal, Bernoulli, Binomial, Poisson, Patterned**, and **Discrete**. The Uniform Distribution is a common choice for it is characterized by lower and upper bounds. Therefore, in a uniform distribution, the variables are drawn with equal probability from all values in the range . The last input required is the **Parameters,** which Excel uses to characterize a distribution. Finally, Excel needs to know where to output the information generated, therefore, mark **New Worksheet Ply** to insert a new ply into the workbook the current ply resides in and pastes the results into Cell A1 of the new ply.

B. SIMULATING DETERMINISTIC BEHAVIOR

The use of Monte Carlo simulation to model a deterministic behavior (e.g., the area under a curve) is illustrated in this section. An example of area under a nonnegative curve is described below.

Example 5.1: Area Under a Curve

Scenario: Find an approximate value to the area under a nonnegative curve. Specifically, suppose $y = f(x)$ is some given continuous function satisfying $0 \le f(x) \le M$ over the closed interval $a \le x \le b$. Here the number M is simply some constant that bounds the function. Therefore, the area is contained within the rectangle of height M and base length $b - a$.

Now, select a point $P(x, y)$ at random from within the rectangular region. This is done by generating two random numbers, x and y, satisfying $a \le x \le b$ and $0 \le y \le M$ and interpreting them as a point P with coordinates x and y. Once $P(x, y)$ is selected, the question is whether or not it lies within the region below the curve; that is, does the y-coordinate satisfy $0 \le y \le f(x)$? It is necessary to count the points that lie within this region as well as the total number of random points generated. Therefore, an approximate value for the area under the curve can be calculated by the following formula:

$$\frac{\text{Area under the curve}}{\text{Area of rectangle}} \approx \frac{\text{Number of points counted below the curve}}{\text{Total number of random points}}$$

Specifically, give the results of a simulation to obtain the area beneath the curve $y = cos\ x$ over the interval $-Pi/2 \leq x \leq Pi/2$, where $0 \leq \cos x \leq 2$.

Using Excel: First, 100 data points will be generated to approximate the value under the curve, $y = cos\ x$. Therefore, go to a new workbook and under the **TOOLS** command, select **Data Analysis**. Then, scroll down to the **Random Number Generation** tool, and select the **OK** button to enter. The Random Number Generation dialog box automatically appears. Enter **2** for the **Number of Variables** and **100** for the **Number of Random Numbers**. Also, select the **Uniform** choice for **Distribution** and for **Parameters:** enter **0** (zero) for *Between* and **1** for *and*. Finally, for **Output options**, mark **New Worksheet Ply** so the output table is entered on a new sheet. When all appropriate information is entered, select the **OK** button to generate the 100 random numbers uniformly distributed between 0 and 1.

At this point, the use of the Excel Random Number Generation tool is complete. To use this information to approximate the area under $y = cos\ x$, Excel formulas are used to manipulate these newly generated random data points. Use the following steps to complete the modeling process:

Step 1: Generate random numbers for x and another column of random numbers for y using the Excel **Random Number Generation** tool.

Step 2: Scale the generated random numbers to the desired interval, $a \leq x \leq b$, using the following transformation equation:

$$t = b - (b - a)\,x$$

Step 3: Calculate all values $f(x_i)$ for all n random coordinates, $x_1, x_2, ..., x_{100}$.

Step 4: Calculate all differences $y_i - f(x_i)$, $i = 1, ..., 100$.

Step 5: Count the number of 0's and negative values among the differences $y_i - f(x_i)$, and record this number as count.

Step 6: Calculate the area: $Area = M(b - a)(Count) / n$.

The random numbers generated by Excel from Step 1 should be in Columns A and B, Rows 1 through 100, of a new worksheet. Once the numbers are entered on the new worksheet by Excel, headings can be added to better organize the data (see Figure 5.1.1). For Step 2, to scale the random x values, the parameters for x are:

$$-Pi/2 \leq x \leq Pi/2$$

where $a = -Pi/2$ and $b = Pi/2$, so

$$t = Pi/2 - (Pi/2 - (-Pi/2))\,x$$

Then use the following Excel formula in Cell C3 to actually scale the random x numbers:

=PI()/2-(PI()/2-(-PI()/2))*A3

copying the formula down through Cell C102 using the fill handle. The empty set of parentheses after *PI* indicate the number of decimal places is not restricted. Likewise, to scale the random y values , the parameters for y are:

$$0 \le y \le 2$$

where $M = 2$, so

$$t = 2 - (2 - 0)y$$

Then use the following Excel formula in Cell D3 to actually scale the random y numbers:

=2-(2-0)*B3

copying the formula down through Cell D102 using the fill handle (see Figure 5.1.1).

For Step 3, since $y = \cos x$, calculate the new values for $f(x_i)$ from the new scaled values of x using the following Excel formula in Cell E3:

=COS(C3)

Then for Step 4, take the difference of the scaled y and the calculated $f(x_i)$ using the following Excel formula in Cell F3:

=D3-E3

	A	B	C	D	E	F	G
1	Random	Random			Calculated	Difference	Count
2	Number, x	Number, y	Scaled x	Scaled y	$f(x_i)$	$y_i - f(x_i)$	True/False
3	0.1318400	0.2033753	1.1566089	1.5932493	0.4024462	1.1908031	FALSE
4	0.4094363	0.5641041	0.2845142	0.8717917	0.9597981	-0.0880064	TRUE
5	0.1774957	0.4170660	1.0131773	1.1658681	0.5291674	0.6367007	FALSE
6	0.0231941	0.0384533	1.4979300	1.9230934	0.0728018	1.8502915	FALSE
7	0.5672170	0.5301675	-0.2111685	0.9396649	0.9777867	-0.0381218	TRUE
8	0.6749168	0.5965148	-0.5495175	0.8069704	0.8527766	-0.0458062	TRUE
9	0.8945891	0.9208045	-1.2396381	0.1583911	0.3251385	-0.1667475	TRUE
10	0.2542192	0.9975585	0.7721432	0.0048830	0.7164171	-0.7115341	TRUE
11	0.0968352	0.6309091	1.2665795	0.7381817	0.2995461	0.4386356	FALSE
12	0.5732597	0.4912259	-0.2301521	1.0175481	0.9736317	0.0439164	FALSE

Figure 5.1.1: Monte Carlo Simulation Data (partial) for Area Under a Curve
(Remember 100 random numbers were generated for x and y each.)

Step 5 requires an intermediate step. First, to see if the differences, $y_i - f(x_i)$, meets the criteria for 0's and negative numbers, allow Excel to test each difference using the following Excel formula in Cell G3:

=F3<=0

which says that the value in Column F is negative (≤ 0) if its corresponding cell in Column G is TRUE. Be sure to remember to drag all of the formulas down through the range of the random numbers (e.g., through Row 102).

Now, to count the number of 0's and negative values among the differences, $y_i - f(x_i)$, use the following Excel formula which can be placed in any particular cell:

=COUNTIF(G3:G102, TRUE)

This counts all the TRUE values in that given range of cells. The answer is 30. Finally, to calculate the area (Step 6), enter the following Excel formula in any particular cell, or work it out on a calculator:

=2*(PI()/2-(-PI()/2))*H3/100

where H3 is the cell used to count the number of TRUE values and 100 is the number of random points generated. The predicted area for these 100 random numbers is 1.884956, where the true value is 2.0.

To calculate a larger number of trials, which will reduce the deviation between the predicted and true values, just enter a larger range for the **Number of Random Numbers** in the Random Number Generation dialog box, and repeat the Steps 2-6.

Example 5.2: Volume Under a Surface

Scenario: Let us consider finding that part of the volume of the sphere

$$x^2 + y^2 + z^2 \leq 1$$

that lies in the first octant, $x > 0$, $y > 0$, and $z > 0$. The methodology to approximate the volume is similar to that of finding the area under a curve. However, to approximate the volume under the surface, use the following rule:

$$\frac{\text{Volume under surface}}{\text{Volume of Box}} \approx \frac{\text{Number of points counted below surface in 1st octant}}{\text{Total number of points}}$$

Specifically, give the results of a simulation to obtain the volume beneath the surface enclosed by the equation above, $x^2 + y^2 + z^2 \leq 1$, that satisfy $0 \leq x_i \leq 1$, $0 \leq y_i \leq 1$, and $0 \leq z_i \leq 1$. (In general, $a \leq x_i \leq b$, $c \leq y_i \leq d$, and $0 \leq z_i \leq M$.)

Using Excel: Again, let us start with 100 data points to approximate the value under the curve. Therefore, go to a new workbook and under the **TOOLS** command, select **Data Analysis**. Then, scroll down to the **Random Number Generation** tool, and select the **OK** button to enter. The Random Number Generation dialog box automatically appears. Enter **3** (three: x, y, z) for the **Number of Variables** and **100** for the **Number of Random Numbers**. Also, select the **Uniform** choice for **Distribution** and for **Parameters:** enter **0** (zero) for **Between** and **1** for **and**. Finally, for **Output options**, mark **New Worksheet Ply** so the output table is entered on a new sheet. When all appropriate information is entered, select the **OK** button to generate the random numbers.

At this point, the use of the Excel Random Number Generation tool is complete. To finish solving this model, Excel formulas are used to manipulate these newly generated random data points. Once the random numbers are generated, use the following step process to complete the model:

Step 1: Generate random numbers for x and another column of random numbers for y using the Excel **Random Number Generation** tool.

Step 2: Calculate all values $f(x_i, y_i)$ for all random coordinates (x_i, y_i).
(Note: No scaling is required since the parameters were entered between 0 and 1.)

Step 3: Count the number of values where $z_i \leq f(x_i, y_i)$.

Step 4: Calculate the volume: $Volume = M(b - a)(d - c)(Count) / n$, where M is the upper bound for z.

Therefore, the random numbers generated by Excel from Step 1 should be in Columns A and B, Rows 1 through 100, of a new worksheet. Once the numbers are entered on the new worksheet by Excel, headings can be added to better organize the data (see Figure 5.2.1). For Step 2, since no scaling is required, for

$$0 \leq x \leq 1$$
$$0 \leq y \leq 2$$
$$0 \leq z \leq 2$$

therefore, the equation for z is:

$$z = \sqrt{1 - (x^2 + y^2)}$$

Calculate the new values for $f(x_i, y_i)$ from the random values for x and y using the following Excel formula in Cell D3:

=SQRT(1-((A3)^2+(B3)^2))

	A	B	C	D	E
1	Random	Random	Random	Calculated	Count
2	Number, x	Number, y	Number, z	$f(x_i, y_i)$	True/False
3	0.7682730	0.1547899	0.1630909	0.6211254	TRUE
4	0.4092837	0.0327769	0.0502945	0.9118182	TRUE
5	0.9114048	0.3542283	0.1461226	0.2094365	TRUE
6	0.3129978	0.9731437	0.6449171	#NUM!	#NUM!
7	0.7703787	0.1799371	0.4678182	0.6116692	TRUE
8	0.9310282	0.4315317	0.1476791	#NUM!	#NUM!
9	0.1815546	0.5266884	0.6701559	0.8304440	TRUE
10	0.6459243	0.1715751	0.7216712	0.7438709	TRUE
11	0.7328104	0.0877407	0.2400281	0.6747521	TRUE
12	0.4849086	0.0858791	0.4533525	0.8703381	TRUE

Figure 5.2.1: Monte Carlo Simulation Data (partial) for Volume Under a Surface

Step 3 requires an intermediate step. First, to see if $z_i \leq f(x_i, y_i)$, allow Excel to test each inequality using the following Excel formula in Cell E3:

=C3<=D3

which says that the value in Column C is less than or equal to the value in Column D if its corresponding cell in Column E is TRUE. Be sure to remember to drag all of the formulas down through the range of the random numbers (e.g., through Row 102). Now to count the number of TRUE responses, use the following Excel formula, which can be placed in any particular cell:

=COUNTIF(E3:E102, TRUE)

This counts all the TRUE values in that given range of cells. The answer is 51. (Note: In Cell D6 is the Excel symbol #NUM! which indicates that a number is used incorrectly (e.g., a negative number under a square root). This indicator is attributed to computer round-off error. The number is actually 0 (zero) or very close to it, therefore, Cell E6 should read FALSE, which is not added to the COUNT total anyway.).

Finally, to calculate the volume (Step 4), enter the following Excel formula in any particular cell, or work it out on a calculator:

=1*(1-0)*(1-0)*F3/100

where F3 is the cell used to count the number of TRUE values and 100 is the number of random values generated. The predicted volume for these 100 random numbers is 0.5100, where the true value is 0.5236.

To calculate a larger number of trials, which often reduces the deviation between the predicted and true values, just enter a larger range for the **Number of Random Numbers** in the Random Number Generation dialog box. However, because the method is strictly probabilistic, greater numerical accuracy will not necessarily be achieved using a run of 10,000 trials instead of a run of 1,000 trials because the particular run of random numbers may not be that good.

C. SIMULATING PROBABILISTIC BEHAVIOR

One of the keys to good Monte Carlo simulation practices is an understanding of the axioms of probability. The term probability refers to the study of both randomness and uncertainty, as well as the quantifying of the likelihoods associated with various outcomes. In Monte Carlo simulation, sequences of random numbers can be used to simulate a probabilistic experiment. The goal in this section is to show how to model some simple probabilistic behaviors using Excel.

1. How to Use the Excel Histogram

The Histogram tool calculates individual and cumulative frequencies for a cell range of data. It also generates data for the number of occurrences of a value in a data set. For example, in a class of 20 students, if one is interested in the distribution of scores in letter-grade categories, then a histogram table can present the letter-grade boundaries and the number of scores between the lower bound and the current bound. The single most frequent score is called the mode of that cell range of data. In other words, the histogram tool returns the most frequently occurring value in an array or a given range of data.

To locate the **Histogram** tool, on the Menu bar under the **TOOLS** command is the menu selection **Data Analysis**. By scrolling down and selecting the **Data Analysis** tool, a separate window entitled the Data Analysis dialog box appears with a list of analysis tools (the same location the **Regression** and the **Random Number Generation** commands are found). Using the scroll bar, move down until the selection **Histogram** appears (the list appears in alphabetical order). Once the entry **Histogram** has been highlighted, select the **OK** command to exit the window which immediately enters the **Histogram** tool as the **Data Analysis** menu selection. A Histogram dialog box appears which asks for input as well as for what output is to be generated.

The **Histogram** tool calculates individual and cumulative frequencies for a cell range of data. The Histogram dialog box first asks for the **Input Range**, which is the reference for the range of worksheet data to be analyzed. The maximum size of the range is 6400 cells. Data must be numeric; otherwise, Excel displays an error message explaining why the data is entered incorrectly. Then, it asks for the **Bin Range (optional)**, which is an optional set of boundary values that define how Excel will divide and sort the data. These values should be in ascending order. Excel counts the number of data points between the current bin number (e.g., lower bound) and the adjoining higher bin (upper bound), if any. Excel includes the values at the lower bin boundary and excludes the values at the upper bin boundary (e.g., $a \leq x < b$). If the bin range is omitted, Excel creates a set of evenly distributed bins between the data's minimum and maximum values (e.g., Excel divides the distance between the lower bound and the upper bound into equal parts).

Excel then asks where to output the information generated under the **Output options**. Therefore, mark **New Worksheet Ply** to insert a new ply into the workbook in which the current ply resides in; and Excel automatically pastes the results into Cell A1 of the new ply or worksheet. Next, select **Pareto (sorted histogram)** to present the data in order of descending frequency. Also, select **Cumulative Percentage** to generate cumulative percentages. Excel generates an output table column for cumulative percentages and includes a cumulative percentage line in the histogram chart. Finally, mark the **Chart Output** if a histogram chart is desired with the output table.

Example 5.3: A Fair Coin

Scenario: Most people realize that the chance of obtaining a head or a tail on a fair coin is 1/2. What happens if we actually start flipping a coin? Will one out of every two flips be a head?

Probably not. Again, probability is a long-term average. Thus, in the long run, the ratio of heads to the number of flips approaches 0.5. Let us define $f(x)$ as follows, where x is a random number between [0, 1]:

$$f(x) = \begin{cases} \text{Head}, & 0 \le x < 0.5 \\ \text{Tail}, & 0.5 \le x \le 1 \end{cases}$$

Note that $f(x)$ assigns the outcome head or tail with equal likelihood to a number between [0, 1].

Using Excel: Again, let us start with 100 data points to approximate the flip of a fair coin. Use the **Random Number Generation** tool located under **Data Analysis** in the **TOOLS** command. At this point, once the numbers are generated, the use of the Excel Random Number Generation tool is complete. Before using the **Histogram** tool, enter into the bins the intervals specifying how the data is to be counted in Column B (see Figure 5.3.1). Therefore, enter **0** (zero) in Cell B3, **0.5** in Cell B4, and **1.0** in Cell B5, which represents the following:

$$x < 0$$

$$0 \le x < 0.5$$

$$0.5 \le x \le 1.0$$

	A	B
1	Random	Bin
2	Numbers	Ranges
3	0.7631764	0.00
4	0.2407910	0.50
5	0.0274972	1.00
6	0.9454939	
7	0.0462355	
8	0.6129643	
9	0.5065157	
10	0.9425642	
11	0.0904874	
12	0.3600574	

Figure 5.3.1: Random Numbers (partial) and Bins for a Flip of a Fair Coin

Now, to use the **Histogram** tool, under the **TOOLS** command, select **Data Analysis**. Then, scroll down to the **Histogram** tool, and select the **OK** button to enter. The Histogram dialog box automatically appears. Under the **Input Range**, ensure the cursor is in the designated box, then highlight the random numbers (e.g., Cells A3 through A102), and either press the **Tab** button or move the cursor to the next box to enter the data. Remember, a dotted line borders the cells to be entered into the range. Now move the cursor down to the **Bin Range (optional)**, highlight Cells

B3 through B5, and then enter the data in the appropriate manner. Finally, for **Output options**, mark **New Worksheet Ply** (so the output table is entered on a new sheet) and mark both the **Pareto (sorted histogram)** and the **Cumulative Percentage** for additional output. When all appropriate information is entered, select the **OK** button to execute the histogram (see Figure 5.3.2).

	A	B	C
1	*Bin*	*Frequency*	*Cumulative %*
2	0.00	0	.00%
3	0.50	58	58.00%
4	1.00	42	100.00%
5	More	0	100.00%
6			

Figure 5.3.2: Histogram Output for the Flip of a Fair Coin

Interpreting the data, Figure 5.3.2 shows that after 100 flips of a fair coin, 58 times the outcome was heads represented by the bin number 0.5, and 48 times the outcome was tails represented by the bin number 1.00. Column C represents the cumulative occurrences which should add to 100% when all flips are counted. The bin numbers 0.00 and 'More' represent if the coin did something other than land on heads or tails (e.g., landing on its thin edge), which rarely occurs as illustrated by zero frequencies out of 100 flips. As the number of random numbers generated increases, the probability of heads occurring is closer to 50% or half the time.

Example 5.4: Roll of a Fair Die

Scenario: Rolling a fair die adds a new twist to the process. In the flip of a coin, only two events are assigned: heads and tails. Now, one must devise a method to assign six (6) events because a die consists of the numbers {1, 2, 3, 4, 5, 6}. The probability of each event occurring is 1/6 because each number is equally likely to occur.

Using Excel: Start with 100 data points to approximate the roll of a fair die. Use the **Random Number Generation** tool located under **Data Analysis** in the **TOOLS** command. At this point, once the numbers are generated, the use of the Excel Random Number Generation tool is complete. Before using the **Histogram** tool, enter the bin range to specify the intervals for how the data is to be counted, in Column B (see Figure 5.4.1). Therefore, enter **0.1667** in Cell B3, **0.3333** in Cell B4, **0.5** in Cell B5, **0.6667** in Cell B6, **0.8333** in Cell B7, and **1.0** (one) in Cell B8, representing the following intervals:

$$0 \leq x \leq 1/6$$
$$1/6 < x \leq 2/6$$
$$2/6 < x \leq 3/6$$

$3/6 < x \leq 4/6$
$4/6 < x \leq 5/6$
$5/6 < x \leq 1$

Now, to use the **Histogram** tool under the **TOOLS** command, select **Data Analysis**. Then scroll down to the **Histogram** tool and select the **OK** button to enter. The Histogram dialog box automatically appears. Under the **Input Range**, ensure the cursor is in the designated box, and highlight the random numbers (e.g., Cells A3 through A102). Next either press the **Tab** button or move the cursor to the next box to enter the data. Remember, a dotted line borders the cells to be entered into the range. Now move the cursor down to the **Bin Range (optional)**, highlight Cells B3 through B8, and then enter the data for the values of the right endpoints in the subintervals (see Figure 5.4.1).

	A	B
1	Random	Bin
2	Numbers	Range
3	0.1643727	0.1667
4	0.8305307	0.3333
5	0.8869594	0.5000
6	0.4549394	0.6667
7	0.3337199	0.8333
8	0.0657979	1.0000
9	0.9139988	
10	0.2865383	
11	0.7667470	
12	0.9852901	

Figure 5.4.1: Random Numbers (partial) and Bin Range for the Roll of a Fair Die

Finally, for **Output options**, mark **New Worksheet Ply**, and then mark the **Pareto (sorted histogram)** for additional output. When all the appropriate information is entered, select the **OK** button to execute the histogram (see Figure 5.4.2).

	A	B	C
1	Die Value	*Bin*	*Frequency*
2	1	0.1667	22
3	2	0.3333	11
4	3	0.5000	17
5	4	0.6667	11
6	5	0.8333	15
7	6	1.0000	24
8		More	0
9			

Figure 5.4.2: Histogram Output for the Roll of a Fair Die

Interpreting the results from Figure 5.4.2, the bin range is equally divided into six intervals between 0 (zero) and one. The Histogram has assigned each interval to one of the six face values on the die. Then the Histogram counts the number of random numbers that occur within each interval (e.g., between 0 and 0.1667, representing the die value of one (1), occurs 22 times out of 100 rolls, etc.). Therefore, the one and the six have occurred more often in 100 rolls. However, since the probability of each event occurring is equally likely, then with more runs or random numbers generated, the frequency of each die value should be the same or occurs 1/6 or 16.67% of the time.

Example 5.5: Roll of an Unfair Die

Scenario: Let us consider a probability model where each event is not equally likely to occur. Assume the die is biased according to the following empirical probability distribution:

Roll Value	P(roll)
1	0.1
2	0.1
3	0.2
4	0.3
5	0.2
6	0.1

Using Excel: Again, we start with 100 data points to approximate the roll of this unfair die. Use the **Random Number Generation** tool located under **Data Analysis** in the **TOOLS** command. At this point, once the numbers are generated, the use of the Excel Random Number Generation tool is complete. Before using the **Histogram** tool, enter the bin range specifying the intervals for how the data is to be counted, in Column B (see Figure 5.5.1). Therefore, enter **0.10** in Cell B3, **0.20** in Cell B4, **0.40** in Cell B5, **0.70** in Cell B6, **0.90** in Cell B7, and **1.0** (one) in Cell B8, which represents the following:

$$0 \le x \le 0.1$$
$$0.1 < x \le 0.2$$
$$0.2 < x \le 0.4$$
$$0.4 < x \le 0.7$$
$$0.7 < x \le 0.9$$
$$0.9 < x \le 1$$

Notice that the length of each interval matches the P(roll) column in the probability distribution.

To use the **Histogram** tool, under the **TOOLS** command select **Data Analysis**. Then, scroll down to the **Histogram** tool, and select the **OK** button to enter. The Histogram dialog box automatically appears. Under the **Input Range**, ensure the cursor is in the designated box, and highlight the random numbers (e.g., Cells A3 through A102). Next either press the **Tab** button or move the cursor to the next box to enter the data. Now move the cursor down to the **Bin Range**

(optional), highlight Cells B3 through B8, and then enter the data showing the values of the right endpoints in the subintervals.

	A	B
1	Random	Bin
2	Numbers	Range
3	0.4709616	0.10
4	0.4998627	0.20
5	0.5854060	0.40
6	0.8194220	0.70
7	0.1235084	0.90
8	0.2588275	1.00
9	0.0120548	
10	0.8679159	
11	0.6740623	
12	0.4740440	

Figure 5.5.1: Random Numbers (partial) and Bin Range for the Roll of an Unfair Die

Finally, for **Output options**, mark **New Worksheet Ply**, and mark the **Pareto (sorted histogram)** for additional output. When all the appropriate information is entered, select the **OK** button to execute the histogram (see Figure 5.5.2).

	A	B	C
1	Die Value	*Bin*	*Frequency*
2	1	0.10	5
3	2	0.20	18
4	3	0.40	20
5	4	0.70	26
6	5	0.90	21
7	6	1.00	10
8		More	0

Figure 5.5.2: Histogram Output for the Roll of an Unfair Coin

Interpreting the results from Figure 5.5.2, one can immediately see from the frequency column that the die values are not occurring an equal number of times. In fact, the intervals or bin ranges are not equal. Since the random numbers are generated between 0 and 1, the bin ranges coincide with the actual range of the numbers generated. If the numbers are generated uniformly between 0 and 1, then after a larger number of trials, the frequency should appear equal to the length of each interval or the difference between each designated bin range (e.g., for the die value of 4, represented by the bin range of 0.70 with a previous bin range of 0.40, the frequency is 0.3 or 30% of the time). Currently, the die value of 4 is occurring 26 times out of 100 rolls or 26/100 = 0.26 or 26% of the time which is approximately 30%.

As demonstrated earlier, Monte Carlo simulation can be used to approximate a deterministic behavior or a probabilistic one. A principle advantage of Monte Carlo simulation is the relative ease with which it can sometimes be used to approximate complex probabilistic systems. Additionally, Monte Carlo simulation provides performance estimation over a wide range of conditions rather than a restricted range as often required by an analytical model. Furthermore, because a particular submodel can be changed rather easily in a Monte Carlo simulation, there is the potential of conducting a sensitivity analysis. On the negative side, the probabilistic nature of the simulation model limits the conclusions that can be drawn from a particular run unless a sensitivity analysis is conducted. Also, the modeler may be forced to develop a simulation model precisely because it is impossible to obtain data, which makes model validation practically impossible. Finally, even though a simulation correctly estimates which of the various alternatives seems best, it still cannot provide an optimal solution because all the possible alternatives have not been considered. Therefore, considerable judgment is required to determine which alternatives to simulate.

VI. LINEAR PROGRAMMING

An optimization problem is said to be a linear program if it satisfies the following properties:

1. There is a unique objective function.
2. Whenever a decision variable appears in either the objective function or one of the constraint functions, it must appear only as a power term with an exponent of one, possibly multiplied by a constant.
3. No term in the objective function or in any of the constraints can contain products of the decision variables.
4. The coefficients of the decision variables in the objective function and each constraint are constants.
5. The decision variables are permitted to assume fractional as well as integer values.

A linear program has the important property that the points satisfying the constraints form a convex set. A convex set is one in which any two points of the set are joined by a straight line segment all of whose points lie within the set. An extreme point of a convex set is any point in the convex set which does not lie on a segment joining some two other points of the set. When a linear function is optimized on a convex set, the optimal solution(s), if one exists, occurs at an extreme point (or a convex combination of extreme points). In this chapter, the user will see how Excel can solve a linear program.

A. LINEAR PROGRAMMING: THE SIMPLEX METHOD

The Simplex Method incorporates both *optimality* and *feasibility* (see Giordano, Weir, and Fox, op. cit., page 316) tests to find the optimal solution(s) to a linear program (if one exists):

An optimality test shows whether an intersection point corresponds to a value of the objective function better that the best value found so far.

A feasibility test determines whether the proposed intersection point is feasible.

To implement the Simplex Method, first separate the decision and slack variables into two nonoverlapping sets, which are called independent and dependent sets. For the particular linear programs considered here, the original independent set will consist of the decision variables, and the slack variables will belong to the dependent set. The Simplex Method consists of the following steps (see Giordano, Weir, and Fox, op. cit., page 320):

Steps of the Simplex Method

1. Tableau Format: Place the linear program in Tableau Format. Adjoin slack variables as needed to convert inequality constraints to equalities. Remember that all variables are nonnegative. Include the

objective function constraint as the last constraint, including its slack variable *z*.

2. Initial Extreme Point: The Simplex Method begins with a known extreme point, usually the origin (0, 0) will be an extreme point.
3. Optimality Test: Determine whether an adjacent intersection point improves the value of the objective function. If not, the current extreme point is optimal. If an improvement is possible, the optimality test determines which variable currently in the independent set (having value zero) should enter the dependent set and become nonzero.

   Examine the last equation (which corresponds to the objective function). If all its coefficients are nonnegative, then stop; the current extreme point is optimal. Otherwise, some variables have negative coefficients, so choose the variables with the largest (in absolute value) negative coefficient as the new entering variable.
4. Feasibility Test: To find a new intersection point, one of the variables in the dependent set must exit to allow the entering variable from Step 3 to become dependent. The feasibility test determines which current dependent variable to choose for exiting, ensuring feasibility.

   Divide the current right-hand side values by the corresponding coefficient values of the entering variable in each equation. Choose the exiting variable to be the one corresponding to the smallest positive ratio after this division.
5. Pivot: Form a new equivalent system of equations by eliminating the new dependent variable from the equations that do not contain the variable that exited in Step 4. Then set the new independent variables to zero in the new system to find the values of the new dependent variables, thereby determining an intersection point.

   Eliminate the entering variable from all the equations that do not contain the exiting variable. Then assign the value 0 (zero) to the variables in the new independent set (consisting of the exited variable and the variables remaining after the entering variable has left to become dependent). The resulting values give the new extreme point (x_1, x_2) and objective function value *z* for that point.
6. Repeat Steps 3–5 until an optimal extreme point is found.

Geometrically, the Simplex Method proceeds from an initial extreme point to an adjacent extreme point until no adjacent extreme point is more optimal. At that time, the current extreme point is an optimal solution.

B. USE OF THE EXCEL SOLVER

The Excel Solver is a powerful optimization and resource allocation tool. It can be used to find the optimum value for a particular cell by adjusting the values of several cells, or to apply specific limitations to one or more of the values involved in a calculation. To use the Excel Solver with a model, first define the problem that needs

to be solved by identifying a target cell, the changing cells, and the constraints used in the analysis:

The *target cell* (also called the *objective* or *objective function*) is the cell in the worksheet model to be minimized, maximized, or set to a certain value.

The *changing cells* (also called *decision variables*) are cells that affect the value of the target cell. Excel Solver adjusts the values of the changing cells procedurally until a solution is found.

Constraint is a cell value specifying certain limits or target values. Constraints may be applied to both the target cell and to changing cells.

Once these elements are specified, Excel is ready to solve the identified problem. (Other optional parameters can be adjusted controlling the reporting options, precision, and the mathematical approach used to arrive at a solution.) After the problem is defined and the solution process initiated, Solver finds values that satisfy the constraints and produce the desired, optimal value for the target cell. Solver then displays the resulting values on the worksheet. There are three types of optimization problems Solver can analyze: linear, nonlinear, and integer (see Giordano, Weir, and Fox, op. cit., page 300 for a discussion on the distinguishing features).

1. Using the Excel Solver

The **Solver** command is located under the **TOOLS** menu on the Menu bar. To use Solver, open the worksheet model and choose the **Solver** command from the **TOOLS** menu. Once the **Solver** command is highlighted, the Solver Parameters dialog box appears (see Figure 6-B.1.1). To specify the target cell (objective function) in the **Set Target Cell** box, enter the name of the cell to be minimized, maximized, or set to a certain value. This target cell should contain a formula that depends, directly or indirectly, on the changing cells specified in the **By Changing Cells** box. If the target cell does not contain a formula, it is considered to be a changing cell. If a target cell is not specified, then Solver seeks a solution satisfying all of the constraints.

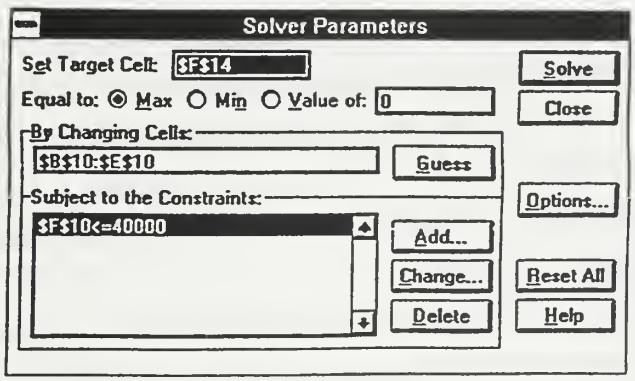


Figure 6-B.1.1: Example of the Solver Parameters Dialog Box

Changing cells normally contain the key parameters of the model. In the **By Changing Cells** box enter the references (or names of the cells) to be changed by Solver until the constraints in the problem are satisfied and

the target cell reaches its goal. If Solver is to propose the changing cells based on the target cell, then choose the **Guess** button. If the **Guess** button is used, a target cell must be specified. Up to 200 changing cells can be specified. Entries in the **By Changing Cells** box usually consist of a reference to a range of cells, or references of several nonadjacent cells separated by commas. (Note: If the changing cells contain formulas, Solver will replace them with constant values assuming the solution is kept.)

In the **Subject To The Constraints** box, build the list of constraints using the **Add**, **Change**, and **Delete** buttons in the Solver Parameters dialog box. Constraints can include upper and lower bounds for any cell in the model, including the target cell and the changing cells. The cell referred to in the **Cell Reference** box under the Add Constraint dialog box (see Figure 6-B.1.2) usually contains a formula that depends, directly or indirectly, on one or more of the cells specified as changing cells. When the Integer (int) operator is used, the constrained value is limited to integers only. (Use integer constraints when a value used in the problem, or the result, must be yes or no (1 or 0), or when decimal values are not desired.) Keep in mind that using the integer method can greatly increase the time necessary for Solver to reach a solution. When using integer constraints, the **Tolerance** setting in the Solver Options dialog box can be used to adjust the allowable margin of error. Only changing cells can be restricted to integer values. For each problem, two constraints for each changing cell (one upper-limit and one lower-limit constraint) can be specified, including up to 100 additional constraints.

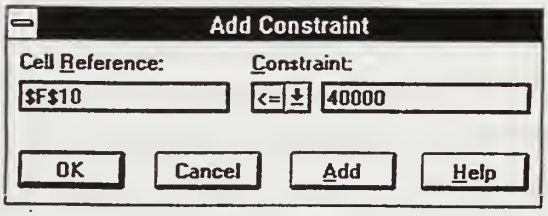


Figure 6-B.1.2: Example of the Add Constraints Dialog Box

Choosing the **Solve** button initiates the problem-solving process. The Excel Solver's solution process involves successive trials, or *iterations*. During each iteration, Solver uses a new set of changing cell values to recalculate the worksheet, and examines the constraints and optimum cell values. The process stops when a solution is found within acceptable precision, no further progress is possible, the maximum time allowed is exceeded, or the maximum number of iterations is reached. When the problem-solving process ends, a dialog box displays several choices:

- Keep the solution Solver found, or restore the original values in the worksheet.
- Save the solution as a named scenario using the **Scenario Manager**.
- View any of Solver's built-in reports.

The model's settings (cell selections, constraints, and options) can be saved by choosing the **Save Model** button in the Solver Options dialog box. When a model is saved, Excel stores its settings as a range of cells containing formulas. (Note: Naming each model range that is saved makes it easier to remember the models in

order to load them later.) To replace the current Solver settings with a set of saved settings, choose the **Load Model** button and enter the name or reference of the range containing the settings to be loaded. Finally, any settings used during a current Excel session remain in effect, but they can be easily reset using the **Reset All** button in the Solver Parameters dialog box. Let us consider a linear programming example.

Example 6.1: The Carpenter's Problem

Scenario: A carpenter makes tables and bookcases for a net profit he estimates as $25 and $30, respectively. He is trying to determine how many of each piece of furniture to make each week. He has up to 690 board feet of lumber to devote weekly to the project and up to 120 hours of labor. He can use the lumber and labor productively elsewhere if they are not used in the production of tables and bookcases. He estimates that it requires 20 board feet of lumber and 5 hours of labor to complete a table and 30 board feet of lumber and 4 hours of labor for a bookcase. He also estimates that he can sell all the tables and bookcases produced. The carpenter wishes to determine a weekly production schedule for tables and bookcases that maximizes his profits.

Let x_1 denote the number of tables to be produced weekly, and let x_2 denote the number of bookcases. Then the model becomes

$$\text{Maximize} \quad 25x_1 + 30x_2$$

Subject to

$$\begin{aligned} 20x_1 + 30x_2 &\le 690 \quad \text{(lumber)} \\ 5x_1 + 4x_2 &\le 120 \quad \text{(labor)} \\ x_1, x_2 &\ge 0 \quad \text{(nonnegativity)} \end{aligned}$$

This model is an example of a linear program. Remember that $25x_1 + 30x_2$ is called the objective function, and x_1 and x_2 are the decision variables. Our attention is restricted to linear programs involving two decision variables.

Using Excel: Excel's Solver tool simplifies the Simplex Method and computes all the calculations for you. All that is required is to accurately enter the target cell, changing cells, and the constraints. Solver does the rest. First, let us enter the carpenter's problem as it is given above (e.g., not in tableau format). Therefore, go to a new worksheet and designate Cells A1 and A2 as the two decision variables, x_1 and x_2. Actually, nothing is typed in those designated cells, but the cells will be referenced in the formulas.

In Cell B1, enter the actual carpenter's problem as an equation using the following Excel formula:

=25*A1+30*A2

Now, enter the first carpenter's constraint equation (without the inequality constraint for it will be entered at a later point) in Cell B3 using the following Excel formula:

=20*A1+30*A2

Enter the second carpenter's constraint equation (without the inequality constraint) in Cell B4 using the following Excel formula:

=5*A1+4*A2

The worksheet should appear similar to Figure 6.1.1. The worksheet is now set up to use the Excel **Solver** tool.

	A	B	C
1		=25*A1+30*A2	
2			
3		=20*A1+30*A2	
4		=5*A1+4*A2	
5			

Figure 6.1.1: Worksheet Preparation for the Carpenter's Problem

For Excel to solve this optimization problem, under the **TOOLS** menu, select the **Solver** command to open the Solver Parameters dialog box. The **Set Target Cell** box is where the equation to be maximized (or the objective function) is entered. Therefore, be sure the cursor is in the **Set Target Cell** box and go to Cell B1. Excel enters the cell reference, **$B$1**, in this box. Next, since this function is to be maximized, be sure the **Equal to** is marked **Max** for maximize. Move the cursor down to the **By Changing Cells** box and go to the first decision variable cell, Cell A1, or type the cell reference for Cell A1, **$A$1**, in this box. To enter the second decision variable cell, insert a comma (,) after the first cell reference and then enter the cell reference, **$A$2**, for Cell A2. Therefore, in the **By Changing Cells** box, the following should appear:

$A$1, $A$2

Now move the cursor to the **Subject to the Constraints** box, and select or highlight the **Add** button. The Add Constraints dialog box should appear. Two entries are required before the constraint is added to the program. The first is a **Cell Reference**, and the other is a numerical **Constraint** which includes a comparison constraint (e.g., <= , = , >= , or **int** for integer). Therefore, to add the first constraint, go to Cell A1 to enter the cell reference **$A$1** in the **Cell Reference** box, then select >= and enter **0** (zero) in the **Constraint** box (since $x_j \geq 0$). Once complete, select the **OK** button to enter the data. Repeat this process until all constraints have been added. The following is a list of all the constraints required to be entered for the carpenter's problem:

Equations or Variables	Cell Reference	Comparison	Constraint
x_1 (decision variable)	A1: **$A$1**	>=	**0**
x_2 (decision variable)	A2: **$A$2**	>=	**0**
$20\,x_1 + 30\,x_2$	B3: **$B$3**	<=	**690**
$5\,x_1 + 4\,x_2$	B4: **$B$4**	<=	**120**

Once these constraints are entered, return to the Solver Parameters dialog box and select the **Solve** button to execute the Excel Solver tool. When the optimization problem is solved, a Solver Results dialog box appears. The dialog box annotates if a solution was found. To save the results of the Excel Solver tool, ensure the **Keep Solver Solution** is marked, and then select the **OK** button which closes and saves the Solver solutions. Before exiting the Solver Results dialog box, the Solver offers three **Report** options: Answer, Sensitivity, and Limits. Right now, just select the Answer Report which automatically inserts a new worksheet entitled *Answer Report 1* and gives a final overview of the problem to include the solution (see Figure 6.1.2).

Target Cell (Max)

Cell	Name	Original Value	Final Value
$B$1		0	750

Adjustable Cells

Cell	Name	Original Value	Final Value
$A$1		0	12
$A$2		0	15

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$B$3		690	$B$3<=690	Binding	0
$B$4		120	$B$4<=120	Binding	0
$A$1		12	$A$1>=0	Not Binding	12
$A$2		15	$A$2>=0	Not Binding	15

Figure 6.1.2: Answer Report for the Carpenter's Problem

Also, the decision variable cells are updated as well as the equation cells with the current solutions in the original worksheet (see Figure 6.1.3). Therefore, Excel found the solution to be:

$x_1 = 12$ (Cell A1)

$x_2 = 15$ (Cell A2)

Maximum: 750 (Cell B1)

	A	B	C
1	12	750	
2	15		
3		690	
4		120	
5			

Figure 6.1.3: Original Worksheet with Updated Results for the Carpenter's Problem

From the results obtained from either the Answer Report generated by the Excel Solver or from the original worksheet input with new updated results, the Excel Solver has concluded that the carpenter should make 12 tables and 15 bookcases to maximize his profit at $750.00 a week.

Example 6.2: The Carpenter's Problem Revisited

Scenario: Since many formats exist for implementing the Simplex Method, let us model the problem in the tableau format and then execute the Excel Solver tool. Again, let x_1 denote the number of tables to be produced weekly, and let x_2 denote the number of bookcases. Then the model is

$$\text{Maximize } 25x_1 + 30x_2$$

Subject to

$$20x_1 + 30x_2 \leq 690 \quad \text{(lumber)}$$
$$5x_1 + 4x_2 \leq 120 \quad \text{(labor)}$$
$$x_1, x_2 \geq 0 \quad \text{(nonnegativity)}$$

Now adjoin a new constraint to ensure that any solution improves the best value of the objective function found thus far. Therefore, the new constraint becomes

$$-25x_1 - 30x_2 \leq 0$$

Next, convert each inequality to an equality by adding a *nonnegative* new variable y_i (or z), called a *slack variable* because it measures the slack or degree of satisfaction of the constant. A negative value for y_i indicates the constraint is not satisfied. (The variable z for the objective function is used to avoid confusion with the other slack variables.) This process gives the augmented constraint set:

$$20x_1 + 30x_2 + y_1 = 690$$
$$5x_1 + 4x_2 + y_2 = 120$$
$$-25x_1 - 30x_2 + z = 0$$

where the variables x_1, x_2, y_1, and y_2 are nonnegative. The value of the variable z represents the value of the objective function. Remember that $25x_1 + 30x_2$ is called the objective function, and y_1, y_2, and z are new slack variables.

Using Excel: Excel's Solver tool executes the Simplex Method and makes all the calculations. All that is required is to enter correctly the target cell, changing cells, and the constraints. First, we enter the carpenter's problem as given above (e.g., in tableau format). Therefore, go to a new worksheet and designate Cells A1 and A2 as the two decision variables, x_1 and x_2, Cells B1 and B2 as the two slack variables y_1 and y_2, and Cell C1 as the slack variable z. Actually, nothing is typed in those designated cells, but the cells will be referenced in the formulas.

In Cell D1, enter the actual carpenter's problem as an equation using the following Excel formula:

=25*A1+30*A2

Now, enter the first carpenter's constraint equation (without the inequality constraint for it will be entered at a later point) in Cell D3 using the following Excel formula:

=20*A1+30*A2+B1

Enter the second carpenter's constraint equation (without the inequality constraint) in Cell D4 using the following Excel formula:

=5*A1+4*A2+B2

Enter the third carpenter's constraint equation (without the inequality constraint) in Cell D5 using the following Excel formula:

=-25*A1-30*A2+C1

The worksheet should appear similar to Figure 6.2.1. The worksheet is now all set to use the Excel **Solver** tool.

	A	B	C	D
1				=25*A1+30*A2
2				
3				=20*A1+30*A2+B1
4				=5*A1+4*A2+B2
5				=-25*A1-30*A2+C1

Figure 6.2.1: Worksheet Preparation for the Carpenter's Problem

For Excel to solve this linear program, under the **TOOLS** menu, select the **Solver** command to open the Solver Parameters dialog box. In the **Set Target Cell** box, this is where the equation to be maximized (or the objective function) is entered. Therefore, be sure the cursor is in the **Set Target Cell** box and go to Cell D1. Excel enters the cell reference, **$D$1**, in this box. Next, since this equation is to be maximized, be sure the **Equal to** is marked **Max**. Move the cursor down to the **By Changing Cells** box and go to the first decision variable cell, Cell A1, or type the cell reference **$A$1** in this box. To enter the second decision variable cell, insert a comma (,) after the first cell reference and then enter the cell reference **$A$2** for Cell A2, as before. Repeat this

process for entering the remaining three slack variables. Therefore, in the **By Changing Cells** box, the following should appear:

$A$1, $A$2, $B$1, $B$2, $C$1

Now move the cursor to the **Subject to the Constraints** box, and select or highlight the **Add** button. The Add Constraints dialog box should appear. Two entries are required before the constraint is added to the program. The first is a **Cell Reference**, and the other is a numerical **Constraint** which includes a comparison constraint (e.g., <= , = , >= , or **int** for integer). Therefore, to add the first constraint, go to Cell A1 to enter the cell reference **$A$1** in the **Cell Reference** box, then select >= and enter **0** (zero) in the **Constraint** box. Once complete, select the **OK** button to enter the data. Repeat this process until all constraints have been added. The following is a list of all the constraints required to be entered for the carpenter's problem:

Equations or Variables	Cell Reference	Comparison	Constraint
x_1 (decision variable)	A1: **$A$1**	>=	**0**
x_2 (decision variable)	A2: **$A$2**	>=	**0**
y_1 (decision variable)	B1: **$B$1**	>=	**0**
y_2 (decision variable)	B2: **$B$2**	>=	**0**
z (decision variable)	C1: **$C$1**	>=	**0**
$20x_1 + 30x_2 + y_1$	D3: **$D$3**	=	**690**
$5x_1 + 4x_2 + y_1$	D4: **$D$4**	=	**120**
$-25x_1 - 30x_2 + z$	D5: **$D$5**	=	**0**

Once these constraints are entered, return to the Solver Parameters dialog box and select the **Solve** button to execute the Excel Solver tool. When the optimization problem is solved, a Solver Results dialog box appears. The dialog box annotates if a solution was found. To save the results of the Excel Solver tool, ensure the **Keep Solver Solution** is marked, and then select the **OK** button which closes and saves the Solver solutions. Before exiting the Solver Results dialog box, the Solver offers three **Report** options: Answer, Sensitivity, and Limits. Again, go ahead and select the Answer Report which automatically inserts a new worksheet entitled *Answer Report 1* and gives a final overview of the problem to include the solution (see Figure 6.2.3).

Also, the decision variable cells are updated as well as the equation cells with the current solutions in the original worksheet (see Figure 6.2.2). Therefore, Excel found the solution to be:

$x_1 = 12$ (Cell A1)

$x_2 = 15$ (Cell A2)

Maximum: 750 (Cells C1 and D1)

	A	B	C	D	E
1	12.00	0.00	750.00	750.00	
2	15.00	0.00			
3				690.00	
4				120.00	
5				0.00	

Figure 6.2.2: Original Worksheet with Updated Results for the Carpenter's Problem (Tableau Format)

Target Cell (Max)

Cell	Name	Original Value	Final Value
$D$1		0.00	750.00

Adjustable Cells

Cell	Name	Original Value	Final Value
$A$1		0.00	12.00
$A$2		0.00	15.00
$B$1		0.00	0.00
$B$2		0.00	0.00
$C$1		0.00	750.00

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$D$3		690.00	$D$3=690	Binding	0.00
$D$4		120.00	$D$4=120	Binding	0.00
$D$5		0.00	$D$5=0	Binding	0.00
$A$1		12.00	$A$1>=0	Not Binding	12.00
$A$2		15.00	$A$2>=0	Not Binding	15.00
$B$1		0.00	$B$1>=0	Binding	0.00
$B$2		0.00	$B$2>=0	Binding	0.00
$C$1		750.00	$C$1>=0	Not Binding	750.00

Figure 6.2.3: Answer Report for the Carpenter's Problem (Tableau Format)

From the results obtained from either the Answer Report generated by the Excel Solver or from the original worksheet input with new updated results, the Excel Solver has concluded that the carpenter should make 12 tables and 15 bookcases to maximize his profit at $750.00 a week, which are the exact same results obtained without using the tableau format. Therefore, either format or technique will work.

C. LINEAR PROGRAMMING: SENSITIVITY ANALYSIS

In addition to solving a linear program, it is nice to know how sensitive the optimal solution is to changes in the constants used to formulate the program. The analysis gives the value of a unit of resource in terms of the value of the objective function at the optimum extreme point, which is a *marginal value*. Sensitivity analysis is a powerful methodology for interpreting linear programs. The information embodied in a carefully accomplished

sensitivity analysis is often at least as valuable to the decision maker as the optimal solution to the linear program itself.

1. Using the Solver Sensitivity Report

The Sensitivity Report contains information demonstrating how sensitive a solution is to changes in the formulas used in the problem. There are two versions of this report, depending on whether it is a linear or a nonlinear problem. The Sensitivity Reports for nonlinear problems are: Reduced Gradient and Lagrange Multiplier. The Reduced Gradient measures the increase in the target cell per unit increase in the changing cell. The Lagrange Multiplier measures the increase in the target cell per unit increase of the corresponding constraint. To obtain a Sensitivity Report, when the Solver Results dialog box appears after solving the optimization problem, select the **Sensitivity** option under **Report**, and then select the **OK** button to execute. Excel, therefore, adds a new sheet entitled Sensitivity Report 1 to the current workbook.

The Sensitivity Report for linear problems adds the following information for each changing cell: Reduced Cost, Objective Coefficient, Allowable Increase, and Allowable Decrease. The Reduced Cost replaces the Reduced Gradient, and measures the increase in the target cell per unit increase in the changing cell. The Objective Coefficient measures the relative relationship between a changing cell and the target cell (objective function). The Allowable Increase (Decrease) shows the change in the objective coefficient before there would be an increase (decrease) in the optimal value of any of the changing cells.

The Sensitivity Report for linear problems also adds the following information for each constraint cell: Shadow Price, Constraint RH Side (right-hand side), Allowable Increase, and Allowable Decrease. The Shadow Price replaces the Lagrange Multiplier, and measures the increase in the objective per unit increase in the right side of the constraint equation. The Constraint RH Side lists the constraint values specified. The allowable Increase (Decrease) shows the change in the Constraint RH Side value before there would be an increase (decrease) in the optimal value of any of the changing cells.

Example 6.3: The Carpenter's Problem Revisited Again

Scenario: Consider again the carpenter's problem. The objective function is to maximize profits where each table nets $25 and each bookcase $30. If z represents the amount of profit, then

$$\text{Maximize } z = 25x_1 + 30x_2$$

Now we ask the question, what is the effect of changing the value of the profit for each table? Intuitively, if we increase the profit sufficiently, we eventually make only tables, instead of the current mix of 12 tables and 15 bookcases. Similarly, if we decrease the profit per table sufficiently, we should make only bookcases. Also, what is the effect of increasing the amount of labor? Currently, there are 120 units of labor available, all of which are used to produce the

12 tables and 15 bookcases represented by the optimal solution. Finally, what does an additional unit of labor worth in terms of the objective function?

Using Excel: After entering all the parameters including the target cell, changing cell(s), and constraints, in t he Solver Parameters dialog box, select the **Options** button before the **Solve** button is highlighted. Since this is a linear program, mark the **Assume Linear Model** in the Solver Options dialog box and then select the **OK** button to return to the Solver parameters dialog box. Marking this option creates a thorough Sensitivity Report including Reduced Cost, Objective Coefficient, Allowable Increase, and Allowable Decrease. Select the **Solve** button to execute the Excel Solver tool. When the linear program is solved, a Solver Results dialog box appears. To obtain a Sensitivity Report, select the **Sensitivity** option under **Report**, and then select the **OK** button to execute. Excel, therefore, automatically adds a new sheet entitled *Sensitivity Report 1* to the current workbook.

For the carpenter's problem outlined in Example 6.1, re-enter the information as before to execute the Solver, but be sure to mark **Assume Linear Model** under the Solver Options dialog box. When the linear program is solved and the Solver Results dialog box appears, select the **Sensitivity** option under **Report**, and then select the **OK** button to execute. Figure 6.3.1 is the Sensitivity Report for the carpenter's problem.

Changing Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$A$1		12.00	0.00	25	12.5	5
$A$2		15.00	0.00	30	7.5	10
$B$1		0.00	-0.71	0	0.714285714	1E+30
$B$2		0.00	-2.14	0	2.142857143	1E+30
$C$1		750.00	0.00	0	1E+30	1

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$D$3		690.00	0.71	690	210	210
$D$4		120.00	2.14	120	52.5	28
$D$5		0.00	0.00	0	1E+30	750

Figure 6.3.1: The Sensitivity Report for the Carpenter's Problem

Interpreting the results from Figure 6.3.1, the first line of the changing cells represents the tables. The carpenter can produce in increase of 12.5 (Allowable Increase) tables or a total of 37.5 (e.g., 25 + 12.5 = 37.5) tables to maintain a profit per table. If the profit per table is reduced by 5 (Allowable Decrease) or goes below 20 (e.g., 25 - 5 = 20) tables, then the carpenter should produce

only bookcases. If the profit per table is between 20 and 37.5 tables, then the carpenter should produce the mix of 12 tables and 15 bookcases. Likewise, the same interpretation can be made of the second line which represents the bookcases.

Interpreting the second line of the constraints represents the units of labor available. The carpenter has only 120 hours of labor to devote to producing the 12 tables and 15 bookcases represented by the optimal solution. The labor constraint can increase by 52.5 hours (or 172.5 hours total) before the carpenter has excess labor. If the labor constraint decreases by 28 hours (Allowable Decrease), then the carpenter has too much lumber and not enough labor.

Finally, the Shadow Price measures the increase in the objective per unit increase in the right hand side of the constraint. Therefore, the report states that the change in the objective function for one unit of change in labor is 2.14 (Shadow Price column). This says that as one more hour of labor is added, the objective function increases by $2.14 as long as the total amount of labor does not exceed 172.5 hours. Thus in terms of the objective function, an additional unit of labor is worth 2.14 units. If management can procure a unit of labor for less than 2.14, it would be profitable to do so. Conversely, if management can sell labor for more than 2.14, it should also consider doing that.

Sensitivity analysis is a powerful methodology for interpreting linear programs. The information embodied in a carefully accomplished sensitivity analysis is often at least as valuable to the decision maker as the optimal solution to the linear program itself.

VII. CONCLUSION

Mathematical modeling is the application of mathematics to explain or predict real world behavior. The primary purpose of this thesis is to provide instructions and examples for using the spreadsheet program Microsoft Excel to support the teaching of a wide range of mathematical modeling applications. Microsoft Excel is a powerful spreadsheet program that allows the user to organize numerical data into an easy-to-follow on-screen grid of columns and rows. In this thesis, it is not the intent to teach mathematical modeling, but to provide computer support for the modeling topics covered in A First Course in Mathematical Modeling by Frank Giordano, Maurice Weir, and William Fox.

This thesis provides an introduction to the entire modeling process using Excel. The Excel software is chosen because of its versatility, its ability to update data quickly and efficiently, and its unique and convenient spreadsheet features. However, Excel does have some disadvantages. First of all, even though Excel does have the capability to solve a system of equations (e.g., by applying the Solver tool), it cannot solve all ill-conditioned matrices. In fact, Excel has trouble solving ill-conditioned matrices regardless of the degree of the polynomial. Therefore, in order to fit a high-order polynomial, one has to solve a system of equations to force the polynomial to satisfy each data point. However, instead of solving the system, Excel can regress the data points on the correct order (higher) of the polynomial (e.g., by applying the Regression tool), but Excel is limited to the degree of the high-order polynomial it can regress ($\leq$ degree of 14). Therefore, the method is limited and does not work for every case.

The other disadvantage for Excel is that it cannot calculate cubic splines (e.g., it does not have a cubic spline command). This is definitely a weakness in the software. Since there is no other alternative to solving cubic splines, a code solving cubic splines is written in Maple and entered instead as an option for the user.

Typical applications for which the user will find Excel useful are in graphical displays of data, transforming data, least-squares curve fitting or linear regression, divided difference tables, programming simulation models, and linear programming. In Chapter I, a condensed introduction to Excel is given which can be used later as a reference. Chapter II expands on the current working knowledge presented thus far on Excel and begins to apply this information to modeling discrete dynamical systems. Data is formulated and charts are created from this data. This is where Excel's power is demonstrated. The charts are easy to read as well as to edit.

In Chapter III, Excel's Function Wizard is introduced that simplifies the formulas and facilitates tabulating the data. More plotting of data is conducted in order to test for proportionality. Here, Excel can enter the slope of a fitted curve to a group of data points without the user making any calculations. Chapter IV discusses model fitting as well as empirical model construction. Here, the disadvantage of Excel is realized, and although the Regression tool is used to best fit a given set of data, Excel cannot overcome this flaw. Chapter V introduces the Random Number Generation and the Histogram tools, both of which make simulation modeling quite

elementary. Finally, Chapter VI introduces Excel's Solver tool which solves optimization models. No linear programming is required since Excel makes all the calculations internally.

There are still areas for development in mathematical modeling as well as in probability and statistics using Excel. In fact, Excel has several built-in data analysis tools suitable for a statistics course. This could be an area for future educational research in developing applications to engineering and statistical scenarios. Also, it is recommended that Microsoft Excel invest in a cubic spline function in addition to the tools already in its software package. Overall, the Excel spreadsheet is an excellent tool for conducting mathematical modeling applications, and therefore, this document provides a valuable source for beginner modelers and demonstrates a sound technique for solving practical modeling problems.

APPENDIX A. RETRIEVING AND ENTERING DATA FROM OTHER FILES

Microsoft Excel can open and save files in many different formats from several applications, including Lotus 1-2-3 and Quattro Pro. The user will find that dealing with files in Excel is much the same as dealing with them in other Windows programs. File formats determine the way the information is stored in a file. Different programs use different file formats. Excel classifies files based on their file name extension. Just remember that an Excel file, which has an *XLS* file name extension, is an entire workbook and will be filed under the Microsoft Excel Files. Any file without the *XLS* file name extension, Excel considers not to be an Excel workbook file, and therefore does not save it under the Microsoft Excel Files. Excel will automatically record the file under a specific type of file if the file name extension matches an available application (e.g., Lotus 1-2-3 Files, Quattro Pro/DOS Files, Text Files, etc.). If the file name extension is too general, then Excel will place the file under an All Files listing. If the user needs to find a file from a list of many files and does not know the name of the file, it is sometimes helpful to know what format the file is under in order to reduce the list and save time.

1. Opening a Document Created by Anotber Application

If the user is working in Excel and would like to open a file from another application, then to specify the type of file to open, choose **OPEN** from the **FILE** menu, and then select a file format from the Open dialog box under *List Files of Type* (the list of supported file formats appears here) (See Figure A.1.1). Also, specify the Drive and the Directory for Excel to look in. If the file is still not listed, select the **Find File** button, and if necessary, select **Search** to give more specific information about the file.

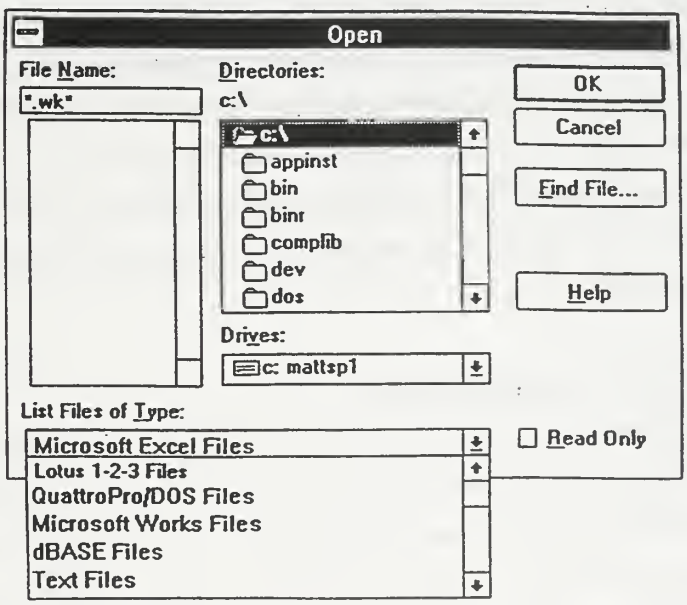


Figure A.1.1: Partial Listing of the Different File Formats in the Open Dialog Box

Opening a file from another application directly is the best technique to transcribe data into Excel. When opening a file in another application's format, Microsoft Excel preserves as much of the original document as possible, including formatting. If Excel cannot accept a formula when opening a file from another application, it substitutes the value calculated by the formula. Although Excel is compatible with some applications, however, if the other format is more current than the Microsoft Excel version, then Excel will not be able to open that file.

When the user is unable to open a file from another application, there are alternatives to retrieving the data into Microsoft Excel. One way is to copy and paste the data from another application directly into a Microsoft Excel document. To accomplish this, the file to be retrieved must be in a Windows-based program (e.g., Quattro Pro); then the user can enter the program, open the document containing the data to be transferred, and highlight the information to be copied. Next, go into that program's **Edit** menu and select **Copy**. (Copying procedures can vary depending on the application. For more information about copying, see that application's documentation.) Switch to the Excel document where the copied data will be pasted, and select the upper-left cell of the area to paste the data. From the **EDIT** menu or the Shortcut Menu, choose **Paste Special**, and select the format in which to paste the data. Finally, select the **OK** button. If the **Paste** command is selected instead of the **Paste Special** command, then all formulas and graphs will not be transferred over and only data and values calculated by formulas will be copied. The **Paste Special** command allows formulas and graphs to be copied to Excel from another application. However, the formulas cannot be edited or seen in the formula bar unless the user double clicks onto a particular cell causing another edit window to open. It is here that the formulas can now be seen and edited. This technique of copying is also referred to as *Embedding* which means inserting information.

One other method that allows the user to transfer data from another application into Microsoft Excel is to drag data between applications. However, this is the least preferred method because the formulas and graphs cannot be dragged or moved between applications. To drag data from another application into Microsoft Excel, arrange both application windows so that the source document and the destination Microsoft Excel document can both be seen on the screen. In the source document, select the data to be transferred, then drag the data to the destination document. Release the data at the location for the data to be inserted. The data now appears in the Microsoft Excel document. Even though this is the least preferred method, it might be the only available way to retrieve the necessary data into Excel.

2. Saving a Document in a Different File Format

If the user is working in Excel and would like to save a file, then to specify the type of file to save, choose **SAVE AS** from the **FILE** menu, and then select a file format from the Save As dialog box under *Save File as Type* (the list of supported file formats appears here) (See Figure A.2.1). Also, remember to specify the Drive and the Directory for Excel to save the file in. When a file is saved from another application after working on it in Microsoft Excel, be sure to specify which file format to save it in. Normally, Microsoft Excel saves a file in

the same format the file was stored in when the file was opened. In particular, if the user is working on a file from Lotus 1-2-3 and used features that are available only in Excel, then the user must explicitly specify that the file is to be saved in Microsoft Excel format with an *XLS* file name extension. By default, Excel saves this file in Lotus 1-2-3 format if another format is not specified, which discards all of the Microsoft Excel-specific work.

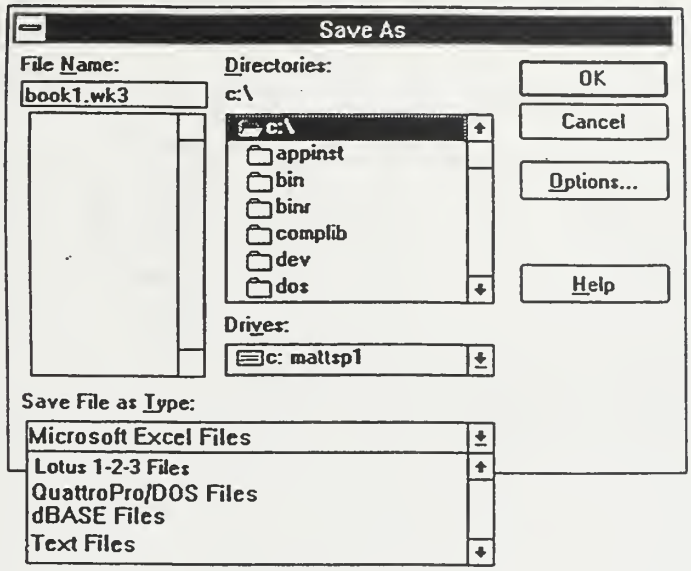


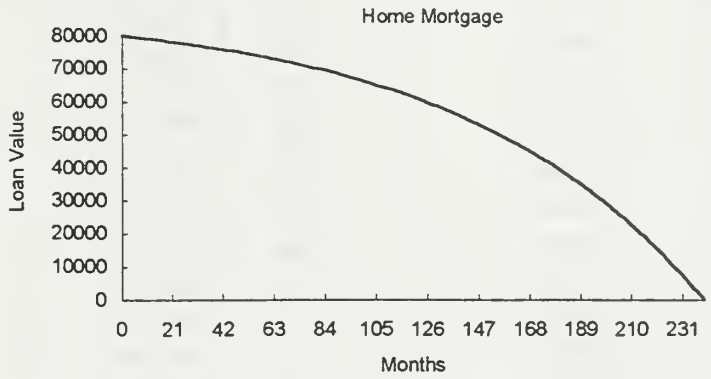
Figure A.2.1: Partial Listing of the Different File Formats in the Save As Dialog Box

Also, files can be opened from earlier versions of Microsoft Excel (e.g., Microsoft Excel 2.1, 3.0, and 4.0), but a document saved in the current version of Microsoft Excel (5.0) cannot be opened in earlier versions of Excel because it uses an updated file format. However, Excel provides the earlier version formats in the selected list of *Save File as Type*, which allows the document to be saved in the format of an earlier version of Microsoft Excel, but the current Excel features will be lost in the resulting file.

APPENDIX B. NUMERICAL DATA FOR MORTGAGING A HOME

The following data were compiled from Example 2.3 in Chapter II and used to produce the associated graph.

Months n	Amount Owed b_n
0	80000.00
1	79919.13
2	79837.45
3	79754.96
4	79671.64
5	79587.48
6	79502.49
7	79416.64
8	79329.94
9	79242.37
10	79153.92
11	79064.59
12	78974.37
13	78883.24
14	78791.20
15	78698.24
16	78604.36
17	78509.53
18	78413.76
19	78317.02
20	78219.32
21	78120.65
22	78020.98
23	77920.32
24	77818.66
25	77715.97
26	77612.26
27	77507.51
28	77401.72
29	77294.87
30	77186.95
31	77077.95
32	76967.85
33	76856.66
34	76744.36
35	76630.93
36	76516.37
37	76400.67
38	76283.80
39	76165.77
40	76046.56



Months n	Amount Owed b_n	Months n	Amount Owed b_n	Months n	Amount Owed b_n
41	75926.15	89	68480.97	137	56477.64
42	75804.55	90	68284.91	138	56161.55
43	75681.72	91	68086.89	139	55842.30
44	75557.67	92	67886.89	140	55519.85
45	75432.38	93	67684.88	141	55194.18
46	75305.83	94	67480.86	142	54865.25
47	75178.02	95	67274.80	143	54533.03
48	75048.93	96	67066.68	144	54197.49
49	74918.55	97	66856.48	145	53858.60
50	74786.86	98	66644.17	146	53516.31
51	74653.86	99	66429.74	147	53170.61
52	74519.53	100	66213.17	148	52821.44
53	74383.85	101	65994.43	149	52468.79
54	74246.82	102	65773.51	150	52112.60
55	74108.42	103	65550.37	151	51752.86
56	73968.64	104	65325.01	152	51389.52
57	73827.45	105	65097.39	153	51022.54
58	73684.86	106	64867.49	154	50651.90
59	73540.84	107	64635.29	155	50277.55
60	73395.37	108	64400.78	156	49899.45
61	73248.46	109	64163.91	157	49517.58
62	73100.07	110	63924.68	158	49131.88
63	72950.20	111	63683.06	159	48742.33
64	72798.83	112	63439.02	160	48348.89
65	72645.95	113	63192.54	161	47951.51
66	72491.54	114	62943.60	162	47550.15
67	72335.59	115	62692.16	163	47144.78
68	72178.07	116	62438.21	164	46735.36
69	72018.98	117	62181.73	165	46321.84
70	71858.30	118	61922.67	166	45904.19
71	71696.02	119	61661.03	167	45482.36
72	71532.11	120	61396.77	168	45056.32
73	71366.56	121	61129.87	169	44626.01
74	71199.35	122	60860.30	170	44191.40
75	71030.48	123	60588.03	171	43752.44
76	70859.91	124	60313.04	172	43309.10
77	70687.64	125	60035.30	173	42861.32
78	70513.65	126	59754.78	174	42409.06
79	70337.91	127	59471.46	175	41952.28
80	70160.42	128	59185.31	176	41490.94
81	69981.16	129	58896.29	177	41024.98
82	69800.10	130	58604.38	178	40554.36
83	69617.23	131	58309.56	179	40079.03
84	69432.53	132	58011.78	180	39598.95
85	69245.99	133	57711.03	181	39114.07
86	69057.58	134	57407.27	182	38624.34
87	68867.28	135	57100.47	183	38129.71
88	68675.09	136	56790.61	184	37630.14

Months n	Amount Owed b_n
185	37125.57
186	36615.96
187	36101.25
188	35581.39
189	35056.33
190	34526.03
191	33990.42
192	33449.45
193	32903.08
194	32351.24
195	31793.88
196	31230.95
197	30662.39
198	30088.14
199	29508.15
200	28922.36
201	28330.72
202	27733.15
203	27129.62
204	26520.04
205	25904.37
206	25282.55
207	24654.50
208	24020.18
209	23379.51
210	22732.43
211	22078.89
212	21418.81
213	20752.12
214	20078.78
215	19398.69
216	18711.81
217	18018.06
218	17317.37
219	16609.67
220	15894.90
221	15172.98
222	14443.84
223	13707.41
224	12963.61
225	12212.38
226	11453.63
227	10687.30
228	9913.30
229	9131.56
230	8342.01
231	7544.56
232	6739.13
233	5925.66
234	5104.04
235	4274.21
236	3436.09
237	2589.58
238	1734.60
239	871.08
240	-1.08

APPENDIX C. NUMERICAL SOLUTIONS TO COMPETITIVE HUNTER MODELS

The following data were compiled for the three cases of the spotted owl versus the hawk, as presented in Figure C.1, in the competitive hunter model from Example 2.6 in Chapter II and were used to produce the associated graphs.

	Owls	Hawks
Case 1	151	199
Case 2	149	201
Case 3	10	10

Figure C.1: The Three Cases (Starting Positions) for Owls vs. Hawks

n	Owls	Hawks
1	151.00	199.00
2	151.15	198.60
3	151.36	198.14
4	151.64	197.60
5	152.01	196.96
6	152.47	196.17
7	153.05	195.20
8	153.79	194.00
9	154.71	192.53
10	155.87	190.72
11	157.31	188.48
12	159.12	185.73
13	161.40	182.34
14	164.25	178.18
15	167.83	173.10
16	172.34	166.93
17	178.04	159.47
18	185.26	150.53
19	194.42	139.91
20	206.11	127.48
21	221.05	113.18
22	240.25	97.09
23	264.97	79.57
24	296.88	61.27
25	338.06	43.27
26	391.05	27.00
27	458.70	13.98
28	544.03	5.35
29	649.92	1.13
30	779.17	0.00

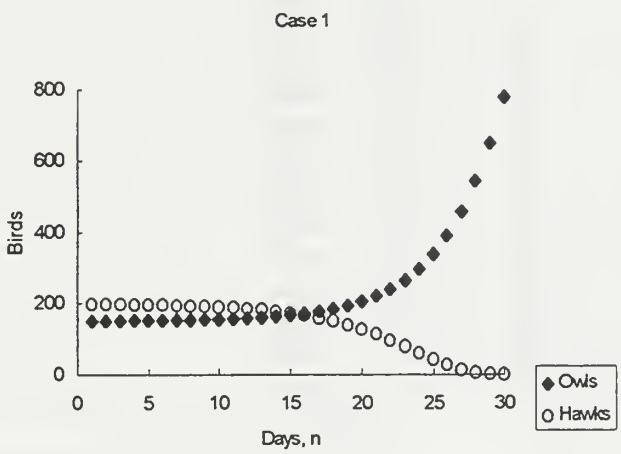


Figure C.2: Case 1, Owls Dominate the Competition

n	Owls	Hawks
1	149.00	201.00
2	148.85	201.40
3	148.64	201.86
4	148.37	202.41
5	148.01	203.07
6	147.55	203.88
7	146.98	204.88
8	146.26	206.12
9	145.37	207.66
10	144.25	209.59
11	142.87	212.00
12	141.16	215.02
13	139.04	218.82
14	136.42	223.62
15	133.20	229.69
16	129.24	237.41
17	124.41	247.27
18	118.53	259.93
19	111.42	276.29
20	102.92	297.61
21	92.88	325.63
22	81.21	362.83
23	67.98	412.75
24	53.52	480.45
25	38.51	573.16
26	24.14	700.97
27	12.05	877.41
28	3.89	1119.50
29	0.31	1446.64
30	-0.08	1879.73

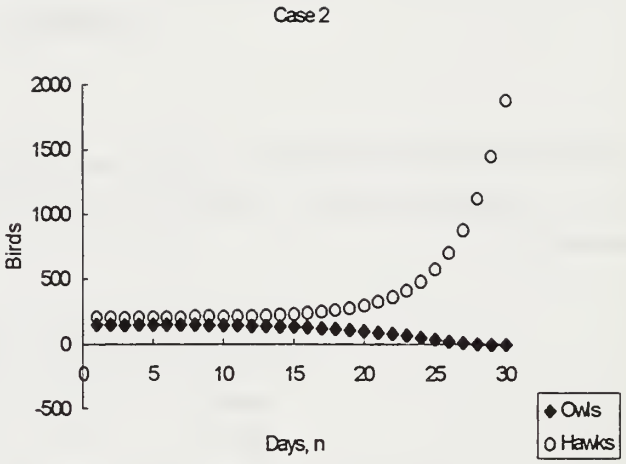


Figure C.3: Case 2, Hawks Dominate the Competition

n	Owls	Hawks
1	10.00	10.00
2	11.90	12.80
3	14.13	16.34
4	16.72	20.77
5	19.72	26.31
6	23.14	33.17
7	27.01	41.58
8	31.28	51.81
9	35.92	64.11
10	40.80	78.74
11	45.75	95.94
12	50.51	115.94
13	54.75	139.01
14	58.09	165.49
15	60.10	195.91
16	60.34	231.14
17	58.47	272.58
18	54.22	322.49
19	47.58	384.26
20	38.81	462.97
21	28.61	565.92
22	18.14	703.32
23	9.01	888.80
24	2.80	1139.43
25	0.17	1474.87
26	-0.05	1916.83
27	0.03	2492.06
28	-0.04	3239.51
29	0.09	4211.65
30	-0.27	5474.40

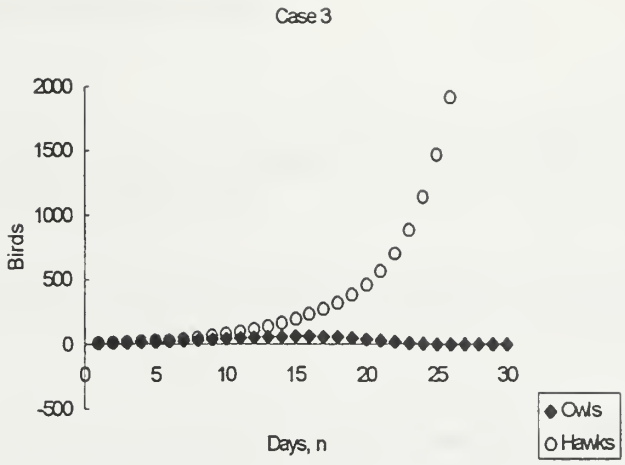


Figure C.4: Case 3, Hawks Dominate the Competition

LIST OF REFERENCES

1. Giordano, F.R., Weir, M.D. , and Fox, W.P., *A First Course in Mathematical Modeling*, 2nd Edition Brooks/Cole Publishing Company, Pacific Grove, California, 1997.

2. Fox, William and et. al., *Mathematical Modeling with Minitab*, Brooks/Cole Publishing Company, Belmont, California, 1987.

3. Microsoft, *User's Guide to Microsoft Excel*, Microsoft Corporation, Cambridge, Massachussetts, 1994.

4. Beauchamp, Robert, *Mathematical Modeling Using Maple*, Naval Postgraduate School, Monterey, California, 1996.

5. Craig, Deborah, *How to Use Microsoft Excel For Windows 95*, Ziff-Davis Press, Emeryville, California, 1995.

6. Weisskopf, Gene, *The ABCs of Excel 97*, Sybex Inc., Alameda, California, 1997.

INITIAL DISTRIBUTION LIST

1. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Road
   Monterey, California 93943-5101

2. Maurice D. Weir, Code MA/Wc 3
   Naval Postgraduate School
   Monterey, California 93943-5216

3. Walter M. Woods, Code MA/Wo 1
   Naval Postgraduate School
   Monterey, California 93943-5216

4. Bard Mansager, Code MA/Ma 1
   Naval Postgraduate School
   Monterey, California 93943-5216

5. Frank R. Giordano 1
   P.O. Box 446
   Seaside, California 93955

6. Nelson L. Emmons, Jr. 3
   3005 Catalina Avenue
   Suffolk, Virginia 23434

7. Defense Technical Information Center..... 2
   8725 John J.Kingman Road, Ste 0944
   Fort Belvoir, VA 22060-6218

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

DUDLEY KNOX LIBRARY
3 2768 00338870 3